

Boycott someone and feel virtuous

Restraints on intellectual traffic between South Africa and the rest of the world have grown over the past year. The academic community should find a better way of making its good intentions come true.

BOYCOTT is an ugly word with no proper place in intellectual life. That is the simplest thing to say about the practice of ostracizing scientists working in South Africa which has become more common since the row, at the end of 1985, about the exclusion of South Africans from the 11th World Archeological Congress eventually at Southampton the following year. The decision to disinvite South Africans was rendered absurd by the fact that they included many who, because of geographical accident (the distribution of ancient australopithecenes, or at least of their fossils), were best placed to contribute to the occasion (see *Nature* 318, 195; 1985). The report from South Africa on page 269 of this issue describes some of the more recent developments in the sanctions business and also, unusually, gives an account of some of the political ramifications of the South African system of apartheid which the boycott is intended to weaken. The hope is that these materials will help readers to form a view of whether the boycott of South African science should be weakened, reinforced or applied in some quite different way.

The most easily won common ground is that apartheid is an abomination whose disappearance would be a boon to everybody, pale-skinned South Africans included. Even liberal South Africans are not always clear why the system gives such great offence elsewhere; in a society long-since accustomed to the use of racial epithets, it is too easily forgotten that there is nothing in human genetics to suggest that a person's racial characteristics can be a reliable predictor of his character, intelligence or other such personal quality. Yet the use of adjectives such as black, white and coloured as labels for describing people is a sign that over-simple valuations of people's virtues are commonplace. Such a system of classification cannot be the basis of a just society. But apartheid is also offensive because, as operated in South Africa, it is the means by which a minority of the population (roughly a quarter of the total now, probably less than a sixth by the end of the century) retains political power in the face of all the odds. Outside South Africa, so much is generally accepted.

The outside world also has a legitimate interest in bringing an end to this state of affairs. Although there is a general presumption that one state has no right to intervene in the internal affairs of another, especially by force, there is nothing to prevent one state, or a group of states, from seeking to influence another. Often, indeed, that is a principal objective of international relations; in the absence of formal international law regulating the nuisance caused by one state to another, winning influence may be the only and the last resort, as illustrated by the continuing efforts of Canada and the Scandinavian states to reduce the flux of acid rain across their borders.

Dependence

There are two particular reasons why South Africa is badly placed to defend the practice of apartheid against the complaints of the world outside. First, South Africa depends on the populations of neighbouring African states for a substantial number of mostly migrant workers who help to keep some of its industries ticking over. The fact that the same workers and the cash they

earn also benefit their home states is beyond dispute, but if anything strengthens the case for seeking influence from outside; neighbouring states may legitimately wish to maximize the benefits while protecting those who earn them from danger and inequity when they are away at work. Second, South Africa appears to have in the past depended on industrial states elsewhere in the world for part of its most skilled labour force, much of it technically trained, which is another legitimate lever of external influence. South Africa would be more immune from external pressure if it were less closely connected with the world outside, but then it would be less prosperous. Even the Botha government does not appear about to reduce uncomfortable pressure by imposing sanctions on itself.

What is called the West has a particular interest in the future of South African society, partly because the white sixth of the population is derived in comparatively recent history from Europe but chiefly because political instability in southern Africa could bring a wider threat. But, South Africans are inclined to say, why pick on us? Are there no other trouble spots and unjust societies? These questions unfortunately beg others. No doubt the governments opposed to apartheid should be more consistent in their willingness to oppose injustice, but can those caught speeding on the roads reasonably ask that their fines should be postponed until all other speeders have been caught? It may also be true that the animus against apartheid is reinforced by unrelated political objectives or even by neighbours' envy of South Africa's prosperity, but that is merely evidence of how easily a position already vulnerable can be made more vulnerable.

Storm

That is why South Africa is for the time being at the centre of a storm of pressure from governments of all kinds. This year, the pressure is likely to increase. The government of Canada, for example, has already announced that it plans to raise the issue of apartheid at the economic summit due later in the northern summer. No doubt the Commonwealth Conference will also return to the theme, taking up the threat to take more direct action against South Africa if the mission of the "eminent persons" should fail (as it did). The later attempt by the British Foreign Secretary to charm success from failure will only have helped increase impatience for change.

So should not a boycott of the South African intellectual community be a part of this process? Unfortunately, too many people and organizations have too quickly jumped to the implied conclusion, even though there is some evidence that the patchy and unpredictable boycott of South African science and its practitioners, and the threat of a more systematic boycott, has already made the technical community in South Africa reflect more carefully about its domestic political environment. So far, too little attention has been paid to the complexities of the intellectual trade, which are much greater even than those of rugby football, too little calculation has been made of what a boycott in science would achieve and almost no thought has been given to the circumstances in which a boycott once imposed would then be lifted.



The analogy of rugby football, to which South Africans are more attached than to science, is instructive. The authorities of rugby-playing nations have agreed not to send representative or 'national' teams to play in South Africa, and not to receive South African teams. From time to time, the rules are broken by unofficial groups, but one consequence of the ban is that apartheid in South African rugby football has partly withered away. Although the process of integration is far from complete, and cannot be while schools (which like to have rugby-football teams) are segregated by race. But from the outset, rugby players and their authorities have been saying that a ban would rob those concerned of the fellow-feeling engendered by competitive sport. But it is also clear that a permanent ban on rugby-playing will not bring the whole apparatus of apartheid crashing down. Logically, it would then make sense if, after a decent interval to ensure that apartheid had vanished from South African rugby, there were to be a relaxation of the ban. It is true, of course, that the Gleneagles agreement which embraces the ban on rugby-playing refers to the severance of all sporting contacts, but that could be renegotiated if rugby-playing could be shown to be a special case.

Science

The boycott of science is much more complicated. The decision of some journals, sometimes at the behest of governments (see p. 273), not to accept for publication articles from South Africa is either a nonsense or stems from the belief that the scientific literature is a sham. By any test, the act of mailing a manuscript to a journal is the gift of a discovery, not necessarily world-shaking, to the international community. A journal's decision to refuse all such manuscripts is to decline, on behalf of the scientific community, what may be a considerable benefit. Especially because South African scientists are productive in several fields, and because South African natural phenomena are often distinctive, the result may be a substantial loss. The argument that publication should nevertheless be banned because the authors acquire prestige thereby falls into the second trap by suggesting that the enhancement of a person's reputation is more than a secondary consideration in publication. In any case, as the world knows, a person denied access to one journal will soon find an outlet in another, perhaps even a South African journal, so that the effect of the boycott is merely to delay the time at which other people know what is new. These are the practical considerations; the principle should be that journals should be open to all those who wish to publish in the field concerned, whether or not they are scoundrels of some kind.

The attendance of South Africans at international conferences is more complicated. Southampton's exclusion of South African palaeoanthropologists last year was absurd because those concerned were bound to have particularly interesting things to say. But, ordinarily, there is no way in which a conference organizer can tell in advance whether unknown applicants for registration will make useful contributions or are merely coming for the ride. Open access by qualified people is therefore as necessary to conferences as to journals. Even those who attend merely to learn, or to be better informed, should not be turned away merely because they live in South Africa, for who is to judge whether the knowledge they gain "will contribute to the ending of apartheid or have no possible role in promoting it"? Physical constraints, such as space limitations, should be determined internationally just as they are at, for example, the Gordon conferences, by peer-review to determine, in advance, those who are likely to contribute most, irrespective of where they come from. There may be a case for closing conferences with a military significance, but South Africans would not usually be alone in this case.

A seductive but insidious variation of this theme is that there should be a selective boycott of those attending conferences. In one version of this proposal, selection would be made by partly professional South African organizations, the newly formed

South African Medical and Dental Association, for example. In another, the African National Congress would certify would-be travellers as being free from racial slur. But this is a recipe for the exercise of prejudice and a precedent for, say, conferences from which supporters of named political parties might be banned.

People's attendance at conferences in South Africa raises greater difficulties. There is a danger that an invited speaker will have such distinction that his or her presence in South Africa will enhance the reputation of the organization concerned. The account of the present balance of opinion in South Africa's universities which appears on p. 274 should satisfy those who are invited by any university that they will not be required to nod apartheid on. The safeguard, otherwise, is to pay for travel and lodging costs independently. And because there is something in the South African complaint that present troubles will hamper the education of the young, people who have the opportunity should jump at the chance of giving groups of lectures at any of the excellent universities in South Africa, of mounting collaborative programmes with them and of encouraging the usual movement of graduate students and postdoctoral fellows.

But will not assistance with the education of young South Africans help to bridge the emerging shortage of skilled people? That argument would be more convincing if it were not that part of the explanation for the present shortage is the inclination of young South Africans to begin their careers abroad. And there is much in the universities' argument that their strength now will determine their capacity to assist when apartheid disappears, peacefully with luck.

Continuing links with South Africa may also help to change the present repressive climate. The danger in further isolation is not so much that South Africans will build a stronger *laager*, but that the one in which they shelter now will never be broken down. Although at least two of the South African universities repeatedly protest publicly at the iniquities of apartheid, and have risked their skins in the process, isolated South Africa's most urgent need is of a credible account of how it may be possible peacefully to effect the transition from the present to a more enlightened society. If voters at the election on 6 May had had a clear picture of how this might be done, what the mathematicians would call an existence theorem, fewer might have been frightened into putting short-term before long-term interests.

Solutions

None of this provides a simple guide as to when and how even this much-attenuated boycott might be ended, but two considerations arise. First, because the ultimate guarantee of equality in the conduct of science in South Africa is that there should be an effective programme of non-white school education, people should wait until such a programme is plainly recognizable. And second, some attempt should be made to match the likely consequences of a boycott of South African science against those of other measures that have or might be taken. If the objective of a science boycott is, for example, to impair the capacity of South African industry to be productive, there are many other ways in which the job could be done effectively. Under the terms of present restrictions by the European Communities, for example, the export of military computers is banned; but effective restrictions on the sale of microchips (not manufactured in South Africa) would quickly have a more immediate and serious effect. Or is it that people are anxious to be seen to be doing something, whatever it is, by personalizing their distaste of apartheid. That would explain the personal rudeness of the past few years, but not how the memory of it will be exorcised when apartheid disappears. If there has to be a boycott, would it not be wiser that it should be polite? In the absence of such or comparable restrictions, it is reasonable to ask why science should be put in the front line. Is it because the cost to the scientific community is literally incalculable (but probably very large), while the loss entailed in refraining from selling microchips is probably smaller, but easily worked out? □

What the British political parties say about science

- Labour and Alliance propose new ministries
- Conservatives promise more of the same

London

As thousands of British scientists unite to press political parties to rescue the country's scientific base from financial crisis, politicians are becoming increasingly aware that, for the first time, science is an election issue. The political parties have all released their policies — some of them sharply contrasting — on research, education, energy and the environment.

The pressure group Save British Science, which represents scientific societies and several thousand individual scientists, including 100 fellows of the Royal Society and 11 Nobel Prize winners, has launched a manifesto of its own. It calls on political parties to plough an additional £100 million a year into top quality research projects, to invest £3,000 million in civil research and development using targeted tax incentives and to establish a long-term strategy by putting scientists in the cabinet and in industry management.

The Labour Party has issued a science manifesto supporting the restoration of research funding as well as a "positive policy" for science. "We need to reduce significantly the proportion of the nation's research and development directed towards military spending and redirect it towards civil priorities as our most successful competitors do", the manifesto states.

Labour aims to reduce the defence share of government research and development from 53 per cent to "about the NATO [North Atlantic Treaty Organisation] average" of 25 per cent. Instead of the present 1.5 per cent of the gross domestic product now spent on civil research and development, Labour would like to see Britain spend 2.5 per cent. This would come from a £400-million increase in the government contribution and a doubling of industry's spending.

Increased spending on research and development is also a feature of the manifesto from the Social Democrat Party/Liberal Alliance. The Alliance also pledges to expand scientific research, and to support the European Communities' joint research programmes. Specific figures for the intended level of investment in science, and their source, are not included.

"A country of our size cannot afford to do everything", the Conservative Party states in the two paragraphs in its manifesto that cover science. It calls for resources to be "better targeted", adding that a Conservative government would

support basic research and "contribute where business cannot realistically be expected to carry all the risks".

In the area of higher education, both Labour and the Alliance are promising to increase the number of student places in universities and polytechnics. Both also say they would provide financial support for more science teachers in schools. Labour specifically pledges to invest in research in higher education institutions.

A re-elected Conservative government would carry through its controversial plans to replace the University Grants Committee with a new University Funding Council made up of academics and industry representatives. Funding would be allocated through a system of contracts. The Conservatives would also abolish local authority control of polytechnics, putting these institutions more under the influence of central government — a move opposed by Labour and the Alliance.

Both opposition parties would abandon plans to build a pressurized water reactor at Sizewell, and both favour energy conservation and the development of renewable energy sources. Labour calls for "gradually diminishing Britain's dependence upon nuclear energy", while the Alliance would push for research into the economics and safety of nuclear power. The Conservatives, on the other hand, promise to develop "abundant, low-cost supplies of nuclear electricity".

The Conservatives say they will press ahead with plans to investigate deep disposal of nuclear waste. Labour and the Alliance both call for further research into disposal strategies.

A new Minister of Environmental Protection would be established under both Labour and the Alliance, and both would develop new pollution control processes. Both also say they would take action to combat acid rain. The Conservatives promise to "continue" to combat acid rain, and pledge to introduce new laws on air pollution and hazardous waste.

The protection of animals used in laboratory experiments is mentioned in all three manifestos. Labour says it would eliminate unnecessary experiments using live animals; the Alliance would set up an Animal Protection Commission to control the welfare of laboratory animals, and the Conservatives point to their record in setting up new legislation on animal experiments.

Kathy Johnston

Hospital workers have AIDS virus

Washington

SHOCK waves have been sent through the medical community by an announcement from the US Centers for Disease Control (CDC) that three health-care workers who were exposed to the blood of patients with AIDS (acquired immune deficiency syndrome) now test positive for human immunodeficiency virus (HIV). The workers' contact with the infected blood did not involve a needleprick or cut, the only ways thought to transfer the virus in a clinical setting.

The announcement, made last week in the CDC's Morbidity and Mortality Weekly Report, described accidents involving only skin-surface exposure to contaminated blood. In the first case, blood seeped through gauze that an emergency room nurse held over the insertion site of an arterial catheter during an attempt to resuscitate a dying patient, reaching her chapped, ungloved hands. Further examination revealed that the patient had died of AIDS and the nurse developed flu-like symptoms 20 days after the incident. Sixteen weeks later, when the nurse donated blood, antibodies to HIV were detected.

The second exposure occurred when a vacuum tube being filled by a phlebotomist with blood from an outpatient suspected of having an HIV infection shattered, splattering her face and mouth with blood. She did not develop symptoms, but tested positive for HIV nine months later. The third worker's ungloved hands and forearms were splashed with blood when a blood-separating machine malfunctioned. She also had flu-like symptoms, and tested positive for HIV after three months.

The workers denied exposure to the HIV virus through sexual relations or intravenous drug use, and none had ever had a blood transfusion. The CDC report suggests that these results demonstrate that "exposure of skin or mucous membranes to contaminated blood may rarely result in transmission of HIV".

In a separate CDC study of 298 health-care workers tested after accidental exposure to HIV-contaminated blood, only one individual was positive for the virus. Eighty-nine per cent of the contacts were through needle-pricks and cuts.

The CDC announcement is expected to tighten adherence to US Public Health Service guidelines for the handling of blood and body fluids from AIDS patients. The guidelines recommend gloves, gowns and goggles to protect medical personnel. But compassionate care is difficult to deliver from behind protective garb, and increased worries about exposure to the virus could further isolate AIDS patients.

Carol Ezzell

Cuts force UK universities into "glamorous" medical research

London

MEDICAL research in British universities has been damaged and distorted by government spending cutbacks in research councils and education, the Committee of Vice-Chancellors and Principals says in evidence prepared for the House of Lord Select Committee on Science and Technology.

Clinical academic staff numbers have dropped by 12.5 per cent since 1981, equivalent to the loss of at least one medical school, according to a quoted survey by the National Association of Health Authorities. "Posts have been downgraded and chairs lost", the vice-chancellors' committee says. "Research has suffered as clinical academic staff understandably put their obligations to patient care and teaching first."

Much of British clinical medical research is carried out in the universities, financed by the University Grants Committee (UGC), the Medical Research Council (MRC), medical charities and foundations, and industry. UGC funding has declined in real terms, and the research councils are also facing financial stringencies; as a result, universities have been relying increasingly on private money. But grants from charities and foundations rarely provide for indirect costs, which must be met from university funds which have already been cut.

"There is a tendency for research to be distorted towards glamorous high-tech projects which attract outside funding from medical charities and industry at the expense of work likely to lead to greater

benefit for the greater number of people", the vice-chancellor's committee says. This "tied" money from the charities and the pharmaceutical industry limits opportunities for innovative, creative research.

The MRC is also increasingly attempting to provide support to specific priority areas, according to the committee; it is no longer able to fund all the alpha-rated research projects submitted to it.

The vice-chancellor's committee's evidence on medical research (like that of the Association of Clinical Professors of Medicine, see *Nature* 327, 180; 1987) was to have been given to the House of Lords last week, but the hearing was cancelled because of the impending election.

Meanwhile, the committee is continuing to fight a battle with the Secretary of State for Education, Mr Kenneth Baker, over the pay of clinical academic staff. The government has allocated £3.25 million for this year's salary award for clinical academic staff, but the vice-chancellors' committee has refused to pay out the money until it receives an assurance that the funds will be built into universities' baseline funding for future years.

Vice-chancellors have threatened that they may not be able to pay the clinical academic staff at the government-decreed level, and may instead seek to negotiate clinical salaries on the basis of universities' ability to pay.

"Universities are no longer willing to subsidize clinical work at the expense of other university work", Mr Maurice Shock, chairman of the committee, told Mr Baker.

Kathy Johnston

Australia's mini-budget putting more pressure on CSIRO

Sydney

IN its May economic statement, known locally as the mini-budget, the Australian government has served notice on the Commonwealth Scientific and Industrial Research Organization (CSIRO) that its budget is to be cut in the next two years. It will lose A\$10 million in 1987-88 and A\$15 million in the following year. Likewise funds for higher education will suffer a cut of A\$12 million in 1987-88 which will grow to A\$24 million in 1988-89.

Even the Australian Research Grants Scheme (ARGS) with its budget of A\$32.1 million has been singled out for a cut of A\$1 million in the next financial year. This comes even though the Minister for Science, Barry Jones, scraped together an extra A\$4.9 million for the ARGS last year to fulfill a longstanding promise that

its budget would grow by 10 per cent in real terms each year.

The government has asked the CSIRO to make up the shortfall in its budget by selling assets and earning more through cooperation with industry. The cut to the higher education sector comes at a time when the universities are suffering from a well-documented decline in their infrastructure.

They are to make ends meet through an "efficiency dividend", an ambiguous term which according to CTEC Research Officer Chris Burgess is taken by the commission to mean that universities should include an element to cover the cost of maintaining infrastructure when charging industry for consulting and other services.

In contrast to the cuts to science and education, the programmes of the Depart-

A moratorium on pork-barrel funds

Washington

AFTER a lengthy and contentious debate, the 56 member institutions of the American Association of Universities (AAU) have agreed, by a vote of about four to one, not to accept 'pork barrel' funds, money which Congress votes directly to particular institutions, bypassing the usual funding agencies and their peer review procedures. The vote is said to be "morally binding", but AAU has no powers of enforcement, and there is sure to be opposition both in Congress and among some non-AAU universities.

The AAU, an organization mostly of well-known research universities including MIT, Yale, Stanford and Cornell, adheres to the philosophy that funds should be allocated strictly according to scientific merit, but this notion has been opposed by smaller universities, which claim, with some historical evidence, that it tends to keep the rich rich and the poor poor. Through the efforts of interested congressmen, a number of large projects have recently been awarded to universities of no great scientific renown (Wichita State and the University of Nevada at Las Vegas are two of the chief beneficiaries, but some AAU members have also profited).

The AAU decision is thus a small tussle in a complex three-way battle. With money hard to obtain, universities differ on the priorities of allocation; at the same time, Congress sees itself as having a responsibility to oversee science funding with more than purely scientific considerations in mind. The AAU made it clear that it did not want its decision to be seen as a challenge to congressional authority, but some of the votes against the moratorium were for precisely that reason.

David Lindley

ment of Industry, Trade and Commerce (DITAC) emerged relatively unscathed. Some areas even received increases. In the opinion of Professor Ron Johnson, director of the Institute for Technology and Social Change at the University of Wollongong, the cuts are another sign to the CSIRO and the universities that the government wants them to develop links with industry more rapidly and to concentrate on more economically relevant research. He believes the government wants an increasing fraction of research supported through the Department of Industry, Trade and Commerce.

One thing seems certain in the light of the May cuts, and that is that the A\$50 million in additional funding for the ARC recommended by ASTEC is unlikely to be forthcoming. As Chris Burgess put it, the chances are "somewhere between zero and none".

Charles Morgan

Scientific breakthroughs regular events on Tokyo stock exchange

Tokyo

WHAT do AIDS (acquired immune deficiency syndrome) and Japan's largest fishing companies have in common? The answer is fish sperm. The link may seem obscure but when brokers on the Tokyo stock exchange recently realized that fish sperm is the raw material of AZT, a possible treatment for AIDS, shares of Nichiro Gyogyo and Nippon Suisan, giant fishing concerns, soared.

With Japan awash in paper money due to a combination of the rising yen, record low interest rates, capital inflow from overseas, and a massive trade surplus, it is boom time on the Tokyo exchange, and the past four months have seen a series of mad upswings in the market riding on the wings of science and technology. First came the AIDS rush in late January after the announcement of the first death of a Japanese woman from the disease (*Nature* 325, 382; 1987). Shares related to AIDS prevention, treatment or testing soared: stocks of Fuji Rebio, a company that has developed the gelatin particle agglutination method for screening blood (*Nature* 323, 384; 1986) doubled in value, and shares of Okamoto, a company reputed to make the world's best condoms, the "skinless skin" made of rubber a mere 100 micrometres thick, quadrupled.

In other cases, the relationship to AIDS was more tenuous. Ueno Seiyaku, a drug maker, announced a new compound that halted AIDS virus proliferation *in vitro*. To the dismay of share dealers, the company was not listed on the stock exchange. But, undeterred, the money-makers latched on to Kyowa Hakko, the world's

largest amino-acid producer, driving up its share price on the basis of a rumour that it would make a licensing agreement with Ueno Seiyaku. In fact, their only relation was that Kyowa Hakko had sold some drugs for the other company.

Now the market has discovered high-temperature superconductivity. First to be singled out were companies listed on a research committee set up by the Science and Technology Agency in February, after Che *et al.*'s announcement of superconductivity in barium-yttrium-copper oxide at 93 K. Prominent among them was Fujikura Densen, an electric cable manufacturer whose shares rocketed from ¥465 in early February to ¥884 on 5 May. As company after company announced "breakthroughs", the market responded — on 2 April, Toshiba's shares jumped ¥40 when the company announced development of wire made from the new ceramics (*Nature* 326, 533; 1987). That the current-carrying capacity of the wire was a mere 6 A cm⁻², five or six orders of magnitude below that for practical applications, did not deter investors.

Such fluctuations are not new to the Tokyo exchange. In early 1985, shares of Asahi Chemical Industry Co. Ltd soared when the company announced the cloning and expression of human tumour necrosis factor (TNF) (*Nature* 313, 803; 1985) but they soon fell back when it became clear that TNF is not a miracle cure for cancer. At the same time, shares of Yamanouchi Pharmaceutical Co. rose on a rumour of a more effective anticancer agent, 'KBS'. To this day nobody knows what KBS was.

David Swinbanks

Star Wars proponents question physicists' beam weapon report

Washington

SUPPORTERS of the Strategic Defense Initiative (SDI) have fired the first round in what may turn into a sniping war over the technical merits of the American Physical Society (APS) report on directed energy weapons (*Nature* 326, 815; 1987). The Science and Engineering Committee for a Secure World, a mostly academic organization devoted to the SDI cause, is publicizing charges made by Gregory Canavan, of Los Alamos National Laboratory, and Lowell Wood, of Lawrence Livermore. They say that the APS study was seriously flawed, containing an accumulation of errors that all make space defence look harder than it really is.

Wood and Canavan's chief point, that the report underestimated the power

achieved by chemical lasers, illustrates the nature of the dispute. The APS panel takes the state of the art to be a laser of 200 kW with adequate beam quality, and says that an increase of power by a factor of one hundred is needed; Wood and Canavan counter that powers of more than a million watts have been achieved but, as they admit, with poorer beam quality. Exactly how much improvement of what is needed for a working weapon?

Arguments such as these show how difficult it is to arrive at definitive answers to ill-defined questions. Because the APS report sets out numerous unsolved technical difficulties but rarely goes so far as to say the problems are insoluble, it leaves room for SDI proponents to call for more research.

David Lindley

Hungary on the ball

London

THE Hungarian Ministry of Industry has established a special research group to investigate possible practical applications of the new superconducting ceramics. The group, headed by Deputy Minister of Industry Arpad Vörös, was set up after a research team at Budapest's Eötvös Lorand University announced the production of a superconducting yttrium-barium-copper oxide ceramic. Last month, the Eötvös Lorand team announced that superconductivity was achieved at 105 K. The initial drop in resistivity was observed at 173 K. The room temperature resistivity of the specimen was very high, at 0.4 ohm cm.

The team, under Dr Jozsef Bankuti, is now investigating the structure of the ceramic using scanning electron microscopy, X-ray diffraction and chemical and microprobe methods. No observations, have so far been made on the Josephson effect, but a Meissner effect of about 21 per cent was observed at 77 K.

V.R.

UK superconductivity

London

MOVES to organize a coordinated strategy between Britain's industry and universities for research into high-temperature superconductors have been initiated by the government's chief scientific adviser, Mr John Fairclough. Representatives of the Cabinet office, the Department of Trade and Industry (DTI) and the Science and Engineering Research Council (SERC) met on 12 May and were due to meet again last week. DTI is expected to meet industrial research directors soon.

Professor Laurie Challis, chairman of the SERC physics committee, says "This is a very hopeful sign which implies that the government is recognizing that this is a special area needing special financial treatment."

S.L.H.

Centred on technology

London

BULGARIA is to establish a network of technological centres to undertake the "technological re-equipping" of the Bulgarian economy. According to Lyubomir Dachev, the deputy chairman of the State Committee for Research and Technology, the centres will concentrate on the many branches of the economy. They will design new equipment, work out plans for the updating of existing capabilities, train specialists and build small high-technology enterprises. Some 25 such centres are planned, in fields including electronics, automation and laser equipment. They will be run by governing councils, Dachev said, and will be entitled to conclude contracts with economic enterprises and to participate in the drawing up of the state economic plan.

V.R.

Bavaria begins mandatory AIDS tests in the face of criticism

Munich

COMPULSORY AIDS tests for prostitutes, intravenous drug users and other suspected AIDS (acquired immune deficiency syndrome) carriers will be introduced in the West German *Land* of Bavaria on 1 June, following a unanimous vote in the Bavarian cabinet. The decision to declare AIDS an epidemic and to take extreme measures to "protect the healthy" has set Bavaria at odds with other *Länder* and provoked strong criticism from the federal government.

The final form of the Bavarian AIDS programme, which has been under discussion since February, retains measures for the "isolation" of AIDS carriers who behave in an "unreasonable" way, such as prostitutes who do not use condoms. But the programme no longer demands that all immigrants from countries outside the European Economic Community be tested for AIDS. Immigrants from other Western European countries such as Austria, Switzerland and Scandinavia will now be exempt from the tests.

Furthermore, people seen "associating with" members of high-risk groups are no longer included in the mandatory testing category. Nor will people accused anonymously be considered for testing. But applicants for public service jobs will still be tested and rejected if found positive for human immunodeficiency virus (HIV).

The measures dealing with the behaviour of prostitutes have been ridiculed along with their author Peter Gauweiler (Christian Social Union), state secretary in the Bavarian Interior Ministry. Prostitutes found not to be AIDS carriers are required to report for quarterly AIDS antibody tests and to use condoms. But when asked how he planned to enforce the latter measure, Gauweiler could only reply that he would depend on "(female) informers from that milieu".

West German officials criticized the Bavarian measures and announced their intention to continue to support publicity and counselling to combat AIDS. West German Health Minister Rita Süßmuth (Christian Democrat) said the Bavarian programme would give the population a "false sense of security". West Berlin Senator Ulf Fink said that the measures would hasten the spread of the disease. There have been about 1,000 cases of AIDS reported in West Germany, and an estimated 50,000 to 100,000 carriers.

In particular, the measures calling for "isolation" of recalcitrant AIDS carriers have met with hostility from a wide range of politicians. Alfons Metzger, speaker for the Bavarian Interior Ministry, defended the measure as a step to "protect

the healthy, not the sick". The hospitals and institutions where violators are to be held have not yet been selected, he said, but that decision will not have to be made "tomorrow morning". The decision to "isolate" someone will proceed through the courts, he explained, "just like the commitment of a dangerous person to a mental institution".

The testing of public service applicants was categorically rejected by the public service union, the *Deutschen Beamten*

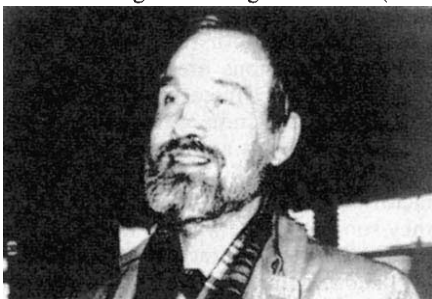
Bund (DBB). DBB speaker Dieter Fengels said that this measure might serve as a signal to private industry and might lead to "stigmatization" of AIDS carriers throughout German society. Metzger defended the measure, comparing it to similar rules forbidding the hiring of diabetics or hypertensives for government service.

Encouraged by the Bavarian cabinet's decision, Gauweiler and his allies in the Christian Social Union, including West German Interior Minister Friedrich Zimmermann, are expected to push for national acceptance of a similar plan. Their chances of winning approval are thought slim. **Steven Dickman**

Warning from Soviet psychiatrist on doubtful practices

London

WESTERN psychiatrists should be wary of Soviet attempts to gain readmission to the World Psychiatric Association, Dr Anatolii Koryagin warned last week. Koryagin, who served a seven-year prison sentence for informing his colleagues abroad (in an



Anatolii Koryagin — beware political psychiatry.

article in *The Lancet*) of the practice of 'political psychiatry' in the Soviet Union, and who was allowed to emigrate to the West a few weeks ago, has been visiting the United States, where he addressed the American Psychiatric Association and also the American Enterprise Institute. Recent reports in the Soviet press of malpractices in Soviet mental hospitals, he warned, are simply an attempt to use the current policy of *glasnost* to detract from the real problem.

The media reports, Koryagin stressed, deal essentially with the practices of corruption, negligence and *laissez-faire* endemic in the Soviet system. The current Gorbachev line encourages criticism of such phenomena in the name of greater efficiency, and so the media are able to criticize psychiatrists who accept bribes from criminals to provide diagnoses which will result in a mitigation of sentence. But the far more serious problem of the use of psychiatric hospitals and neuroleptic drugs as a means of punishing political activists and dissidents remains unchanged.

Western concern about Soviet misuse of psychiatry for political ends led, in 1983, to the tabling of a number of motions of censure to be raised at the Vienna Congress of the World Psychiatric Association (WPA). Before these motions could be formally raised, the Soviet delegation (together with several delegations from the socialist bloc) resigned from the WPA. The Soviet resignation was accepted with "regret", and a motion from the UK Royal College of Psychiatrists by an overwhelming majority assured the Soviets that they would be welcome to rejoin as soon as they can prove they have abandoned doubtful practices. During the past two years, there have been several discreet hints from the Soviet psychiatric establishment that it would like to reapply for admission to the WPA before its next congress, due in Athens in 1987. but, says Koryagin, before the Soviets are readmitted, they should comply, and be seen to comply, with three conditions.

First, he said, the Soviet Union should admit that it has been making improper use of psychiatric methods and drugs to punish its political opponents. This would entail not merely verbal admissions, but also the establishment of some kind of tribunal to attribute responsibility to those concerned in the abuses, and the repudiation of the current Soviet theory that dissent is a symptom of mental disorder.

Second, said Koryagin, all "psychiatric prisoners of conscience" should be released, and, finally, the Soviet All-Union Society of Neuropathologists and Psychiatrists (the professional association for Soviet psychiatrists) should agree to participate in the international commission established by the WPA to investigate reports of psychiatric abuse worldwide — a commission which at present has little more than a paper existence, largely due to the current attitude of the Soviet Union. **Vera Rich**

UK releases of engineered organisms to go ahead

London

THREE of four planned UK experiments involving the deliberate release of genetically manipulated organisms (see *Nature* 326, 537; 1987) are set to go ahead in the coming weeks. All three had been revised in response to comments from the Advisory Committee on Genetic Manipulation (ACGM), which expressed itself happy with the revised plans at a meeting two weeks ago. Revised plans and preparations for the fourth experiment, which involves baculoviruses, have not yet been submitted to ACGM.

Two of the planned experiments involve potatoes, the third involves *Rhizobium*, the nitrogen-fixing bacterium that forms the root nodules of leguminous plants. The potatoes, to be planted in a carefully monitored field near Cambridge by scientists at the Plant Breeding Institute (PBI) of the Agricultural and Food Research Council (AFRC), have been genetically engineered to contain genes for two bacterial enzymes. Whereas neither gene 'improves' the plant, the experiment is a model for future improvements.

Following consultations with ACGM, the PBI scientists have designed the experiment to avoid any spread of the plants from the experimental plot. The plants will be deflowered, weeding and harvesting will be carried out only by hand, the potato tubers will be handled in isolation from others and the field will be kept fallow for a year.

Similar precautions will apply to the potato experiment at the AFRC Institute

of Arable Crops Research at the Rothamsted Experimental Station in Harpenden, Hertfordshire. In this case, the plants derive from cell fusion between a domestic potato and a wild South American species that can resist the potato leaf roll virus. No recombinant DNA technology is involved.

The *Rhizobium* experiment, also at Rothamsted, is designed to test the extent to which genes can be transferred between rhizobial strains in soil. It is financed by a European Economic Community programme for risk assessment in biotechnology, as one of the conceivable hazards of deliberate release experiments is the unwanted transfer of genes between strains. The strain that will be released at Rothamsted contains a harmless marker gene whose transfer to natural strains in the soil will be monitored.

Although the last of these experiments would be expected to meet the approval of Jeremy Rifkin, who was last week exploring the European biotechnology scene, fresh from his defeat in preventing the US field test of ice-minus bacteria (see *Nature* 326, 819; 1987), he was unwilling to be drawn on his views. After three days of pressing his case for a moratorium on all deliberate release experiments "until a proper science of risk assessment can be developed", Rifkin left for West Germany saying only that "It's up to the British" to decide whether to act against UK experiments but admitting that ACGM's deliberations were "distinctly more thoughtful" than in equivalent committees in the United States.

Peter Newmark

Europe-only mobile telephone plan

London

FOUR European governments have signed an accord jointly to develop a digital telephone system, based on cellular radio, which can be used in all states and across national boundaries. The agreement, signed by Britain, West Germany, France and Italy in Bonn last week, is meant to ensure that European nations develop and commercially exploit their own systems to the exclusion of Japanese and US companies which now dominate the technology.

British Telecom (BT) and Racal, which operate a service in the United Kingdom based on related technology, but largely of US origin, also signed the Bonn agreement and have already begun to investigate how the technology can be developed for use across the continent.

The cellular radio system operates by using a range of frequencies. The mobile unit's transmitting and receiving frequen-

cies are automatically changed as it crosses the perimeters of a geographical area or cell. If the system is developed properly, the entire continent would be divided into a honeycomb of cells.

In Britain, two major electronics groups, GEC and Plessey, have already joined forces with BT and Racal to create a prototype for Europe. The £1-million project, unveiled earlier this month, will comply with the standards set by the Conference of European Posts and Telecommunications, the European standards authority. The results of the project are expected by the end of the year.

Most of the European nations have already gained substantial experience in the use of the technology to be adopted in the system but the refinements to the standards need to be finalized by the end of the year, if products are to be ready by the target date of 1991.

Bill Johnstone

New protein database for Europe

Martinsreid

ON-LINE support for researchers using protein sequence data should soon be offered by the Max Planck Institute for Biochemistry at Martinsreid, just outside Munich, West Germany. A new database called MIPS (Martinsreid Institute for Protein Sequence data) will serve as the European partner to the Protein Identification Resource (PIR) offered by the National Biomedical Research Foundation (NBRF) in the United States.

The database will address a key problem of accessibility for European researchers, said project co-leader H. W. Mewes last week. "There is an extremely large demand" for such a service, he said, "but in West Germany, only the ten largest institutions subscribe to existing databases". A computer program offering a database and sequence analysis has proven popular among the scientists at Martinsreid, who have accessed it more than 10,000 times within the past year.

In addition to improving accessibility, the staff of MIPS will also push for speedier processing of protein sequences submitted for publication. The current lag time of 12-18 months between discovery of a sequence and its appearance in a database will be shortened, Mewes hopes, by three months or more.

MIPS has already received support from the European Economic Community (EEC) and is expected to receive a grant of DM 6.6 million within the next six months. Private companies such as Boehringer have also expressed an interest.

The establishment of MIPS should help relieve the pressure on NBRF, which, said Mewes, would have to increase its staff tenfold just to handle the current explosion of data. The MIPS databank will be housed in a minicomputer hooked up to two front-end processors. As with PIR, the data and software used will be released into the public domain without user fees.

Eventually, MIPS will offer several spinoffs from the main sequence databank, including listings of protein fragments and artificial and natural mutants. Both of these aims are already within reach. A long-term goal is to offer a database that reveals the known biological activity and function of a protein when given the structure.

The group at Martinsreid is not the only one interested in improving on the *status quo* in protein databases. Amos Bairoch at the Geneva University Medical Center has created a database called SWISSPROT which is an adaptation of the NBRF database. But Mewes believes that MIPS has shown itself to be the most extensive and promising database project on offer.

Steven Dickman

NASA dreaming of a permanent base on the Moon...

Washington

FIVE contracts for studies of the utility and feasibility of a manned base on the Moon will be awarded this week by the US National Aeronautics and Space Administration (NASA). The contracts are small, in total worth only \$1.2 million a year, and they certainly will not launch any astronauts. But they do give a little substance to the dreams that NASA would like to make reality.

NASA has been having difficult times. The most pressing problems are obvious: space launches have fallen years behind schedule, the shuttle has suffered a tragic accident, and an unwise decision was made to abandon expendable launch vehicles. But the more fundamental concern, as Barney Roberts, Mission Manager for the Advanced Program Office at NASA put it, is that "we are no longer respected. We don't have a space station as the Russians do. The Japanese and Chinese have worked very hard and achieved space capabilities. We are becoming second-rate. We now feel a lot of pressure from Congress, the President, the National Science Council, and the public to get things going."

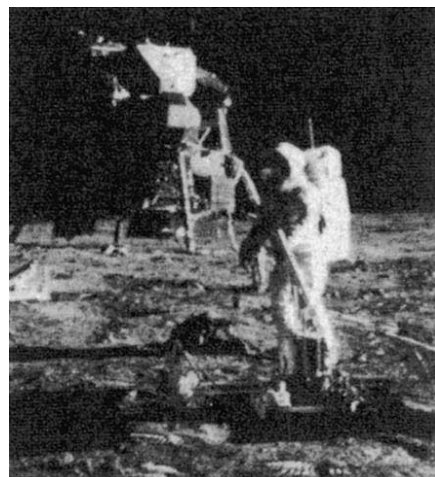
Reports have come out one after another from the National Commission on Space, the Space Science Board of the National Research Council, and several of NASA's own committees. All come to much the same conclusion: NASA must try to return to the heady days of the Apollo programme and have a bold, long-term project that makes US space supremacy obvious to the world. The most influential of the reports, that of the National Commission on Space, whose members were appointed by President Ronald Reagan, declared for a Mars colony last year and celebrated the possibilities of life in space with an artist's impression verging on the psychedelic. More recently, NASA announced that it is studying four initiatives to give the United States a major goal in space; an expanded study of Earth systems, more Solar System exploration, a permanent base on the Moon and human exploration of Mars.

The pressure has been strong to go for Mars; conferences have been organized and Congress lobbied. But there are signs that NASA will leave some of its options open by trying to win support for a Moon-base, that could later provide a step towards Mars. An attractive feature of the Moon is that it could be mined to provide the raw materials for further voyages. Forty per cent of weight of the lunar soil is oxygen. If that could be extracted, rockets could obtain the heaviest component of their fuel in space. There is also plenty of

silicon for solar cells and meteorite bombardment has provided relatively pure iron ore which can be picked up with a magnet.

As a percentage of gross national product, the National Commission on Space has calculated that a lunar base would cost less per year than the Apollo project, although it would run for longer. Man could be back on the Moon by 2000, with a permanent base five to ten years later. The studies commissioned this week are but a tiny step in that direction. Lockheed Corporation will manage two projects to look at uses that could be made of a Moon base, and at new space vehicles to operate in its vicinity. Two further projects will aim at seeing how the different activities of a base could be integrated overall. One has yet to be awarded. The other goes to the University of Texas and the Large Scale Programs Institute, Austin, to develop a computer model of the systems needed to run the base. The final project goes to the University of California, San Diego, to study methods of transport that do not use chemical rockets, including space tethers — miles of wire that could be used to winch spacecraft in and out of orbit.

The idea of a Moon base has already set scientists dreaming — even though most of them would be in their rocking chairs before it could actually be built. An unofficial group centred on the University of New Mexico has been looking at whether the Moon could offer anything unique for



We'll be back? Edward Aldrin deploying a seismic experiments package back in the heady days of the Apollo programme, 1969.

astronomers. It seems that it could. An observatory on the hidden side of the Moon would be sheltered from disturbances produced by the Earth. That means it would be possible to open up astronomy at new, very-low frequencies that auroral kilometric radiation would prevent in an Earth-orbiting observatory.

Optical and infrared interferometry could take advantage of the totally clear lunar atmosphere and the extreme stability of a baseline measured on the Moon's geologically dead surface. With a ten-kilometre baseline, resolution at the middle of the optical region could be as good as one microarcsecond, five times the planned resolution of the Hubble Space Telescope. Use could even be made of huge lunar craters to build 'natural' radio telescopes, like that at Arecibo.

Alun Anderson

... and facing down-to-Earth problems

Washington

RESUMPTION of US space shuttle flights has been delayed yet again. The target date for the first launch since the loss of the Challenger is now June 1988, four months later than the date announced last October.

Even the June date may be optimistic. Two large-scale tests must be performed on the shuttle's main engines, as well as a series of tests on the solid-fuel boosters that caused the 1986 accident. The new plan assumes that all tests will be successful and allows no time for possible mishaps.

Only three launches will now be possible in 1988 and seven in 1989 and there are bound to be further delays for science missions. Waiting in line for an early launch are the Hubble Space Telescope, the ultraviolet astronomy telescope (ASTRO-1), the Venus radar mapper (Magellan), the Jupiter mission (Galileo) and the solar polar mission (Ulysses), along with communications and defence satellites.

Some satellites originally planned for the

shuttle will be launched by the conventional rockets that the shuttle was intended to replace. After a year of rumours, NASA confirmed last week that it will acquire expendable launch vehicles to try, as NASA administrator James Fletcher put it, "to accelerate the deployment of the nation's backlog of space science missions". The West German X-ray satellite ROSAT, the Combined Release and Radiation Effects satellite, and the Extreme Ultraviolet Experiment will be the first to benefit with launches in 1990-91. One of the larger planetary missions may be launched aboard a massive Titan rocket, but a decision whether to do so will not be made for some months.

Early launches will be made by buying rockets tailored to the already designed payloads. Later, competitive bids from industry will be taken in line with the Reagan administration's policy of commercializing the market for expendable launch vehicles.

Alun Anderson

Election further postpones the decision on UK space plans

London

THE long-awaited decision on the British national space plan has been further delayed with the announcement of the general election. The British National Space Centre (BNSC), which last September prepared the 15-year plan at the request of the government, does not now expect a decision until "well into July at the earliest".

The protracted deliberations are causing frustration among BNSC staff and the British space community, who are finding it difficult to formulate long-term research strategies. The speediest solution likely would arise if the Conservative government is returned and if the ministers involved in the original decision-making are retained; if different ministers are appointed, a briefing period, likely to take several weeks, would be necessary.

Given the length of time it is taking for a decision to be made, Britain is fortunate that a ministerial council meeting of the European Space Agency (ESA), originally suggested for next month, is not now likely to be held until November, and BNSC makes no secret of its relief. Roy Gibson, director general of BNSC, says that a decision on the plan is essential for BNSC if it is to retain credibility during negotiations around the ESA table, "particularly as we frequently lead the charge for realism in the costing and time schedules of the new large programmes. What remains of diplomatic courtesy has so far

discouraged our friends in Europe from asking whether the United Kingdom's financial contribution is worth the trouble we are causing, but it has come very close once or twice."

The proposed space plan has not been published and BNSC will not indicate how



much money it has asked for, except that it is a "significant increase" on the present budget of around £110 million, of which 80 per cent goes to ESA. Informed estimates suggest that a figure of around £300 million a year is realistic, given that BNSC is to double its ESA contribution over the next three years and that the plan proposes to increase the proportion of the domestic space programme from 20 per cent of the total spend to nearer 40 per cent.

Simon Hadlington

Ariane's new schedule for keeping ahead

Paris

ALMOST exactly a year after a third-stage ignition failure forced the destruction of the European space rocket, Ariane 2, Arianespace has issued its new launch schedule. The next launch, V19, is now planned for August 1987, if the June trials of the modified third-stage liquid hydrogen-oxygen engine are successful.

If Ariane is given a clean bill of health in June, three launches will be carried out in 1987, eight in 1988 and nine in 1989 and 1990. As Ariane can launch two satellites on one rocket, Arianespace is optimistic that it will be able to make up for lost time. The backlog of orders stands at 46 satellites, worth an estimated \$2,450 million, of which 21 are from European customers, 9 from the United States, 10 from international companies and the remaining 6 from Canada, India, Japan and Australia.

The sudden halt to the Ariane launch schedule last May prevented the European company from taking full advantage of the gap in the market opened by the problems

that had halted shuttle launches. Despite a full order book, Arianespace has had to watch its lead be whittled away as the Soviet Union, China and Japan prepare to enter the commercial market. But Frédéric d'Allest, chief executive officer of Arianespace, feels that Europe has time on its side. In Washington last week d'Allest said that "there is no real threat from China until 1992 at the earliest. The Japanese H-2 rocket is capable of launching medium-sized satellites, but will not be ready until 1994."

The accidents to Ariane 2 and Challenger in 1986 followed a loss to the space insurance market in 1985 exceeding \$400 million — twice the revenue collected from insurance premiums in the previous 13 years. Today, almost half of the cost of launching a class 1/2 Ariane 3 satellite is to cover insurance. In January 1986 Arianespace set up a wholly-owned insurance subsidiary to cover launch-phase risks in an attempt to keep down the total cost to the customer.

Peter Coles

Caesium limits a hot issue

London

NEW limits on radioactivity in food proposed last week by the European Commission could prove a political hot potato. Scientific advice on acceptable caesium levels has been ignored, and the proposed limits seem unlikely to be supported by all member countries.

The European Commission's proposed maximum permissible limits for radioactive contamination of food and drinking water would be introduced as part of a permanent safeguard system to be operated in the event of a nuclear accident, and would replace the emergency limits introduced in the aftermath of the Chernobyl accident. The emergency limits expire on 31 October.

Advice sought from scientists and public health experts who make up the so-called Article 31 Committee of Euratom was accepted for isotopes of plutonium, iodine and strontium, but the Commission sharply reduced the figures for caesium. For caesium the Commission is recommending a limit of 1,000 becquerels per kilogram or litre for dairy products, 1,250 for other foodstuffs, 800 for drinking water, and 2,500 for animal feedstocks. The Article 31 experts had suggested 4,000 for dairy products and 5,000 for other foods.

The Commission explained that some countries outside the European Economic Community (EEC) are applying strict levels that could pose problems for Community exporters. "The levels chosen should command public confidence", a briefing paper explained. But the proposals have been criticized by the director of Britain's National Radiological Protection Board, Dr John Dunster, a member of the Article 31 Committee. "I don't think low levels command public confidence, especially when you muck about with the figures", he said. But, Dunster added, "it is legitimate for scientific advice given to a political body to be rejected on political grounds".

The Article 31 Committee last year made "reasonable guesses" which led to recommended caesium limits of 30,000 becquerels per kilogram, according to Dunster, but the Commission staff cut the level back to 9,000 to allow for additivity. The final post-Chernobyl limits adopted were 600 for dairy products and 370 for other food imported into the EEC. The limits within Britain were set at 1,000.

"The Commission has chickened out", Dunster said. "But if they think 1,000 and 1,250 will be better for trade and public opinion, it's their judgement — but it's not supported by the scientific advice."

Kathy Johnston

The Vatican and embryology

SIR—As a developmental biologist and a Roman Catholic, I would like to make some comments on your leading article "The Vatican and embryology" (*Nature* 326, 229; 1986). The article correctly points out the two principles applied by the Instruction of the Catholic Church: life of individual human beings shall be respected absolutely and sexual acts and procreation must be reciprocally connected. The document implicitly applies another basic principle underlying and justifying the previous two. It is, that man is *more* than an exhaustive addition of experimental data; man is *more* than its functioning. This is a basic issue, pre-ceding the consideration of any research data on man's functioning. One can argue classically that what cannot be studied by science is not scientific, but non-scientific does not mean irrelevant.

Establishing the moment when an individual unborn human being becomes a human person, that it has the same rights as a born human baby, is not a matter of natural sciences, unless one can demonstrate a physio-chemical or biological structure underlying human rights. Most people agree that it is the *more* that provides the criteria for solving the issue. The Catholic Church too agrees to this proposal. And in applying its *more*, considers correctly that its conclusions are not rejected by research data, which have exhaustively shown individuality to be definitely fixed with the fusion of the two haploid gametes. Logically, the Catholic Church also discourages additional research on human (embryonic) persons for the purpose of reaching a better understanding of their functioning for the simple reason that embryos are already people: research comes too late, and its only contribution to that unique human person would probably be its death.

Looking at a similar issue in the past, human slaves did not have to wait for a better understanding of man's functioning to show that they were similar to human non-slaves. Fortunately for them and for us, they were finally recognized to be human persons. Looking at the future, we have to be careful that other old, mentally ill or undesired human beings do not again lose their rights either with the complicity of science or in its name.

Setting no limits for science other than science itself, as the article implicitly claims, is neither scientific, because it is basically a pre-scientific choice, nor acceptable because the only limits would be the scientifically and technically possible. And despite the many marvellous achievements of science, we have enough examples of horrors that science can cause, both under control (Dresden, Hiroshima and Nagasaki, Gulag) and out

of control (thalidomide, Seveso, Spanish oil, Bhopal, Chernobyl). The reasonable limit proposed by the Catholic Church respects what man is.

P. JUNQUERA

CIBA GEIGY,
Centre de recherches agricoles,
1566-St-Aubin, Switzerland

SIR—Your leading article "The Vatican and embryology" (*Nature* 326, 229; 1987) indicates that the writer does not (or does not want to) understand people who believe in the eternal life of every human individual.

The resurrection of Christ is the basis, in reality the only basis, of the Christian faith. If there is no eternal life, scientists as well as militarists, industrial managers or political powers may do with human beings what they want. No man-made agreement will be able to establish common human standards of conduct. Future generations may have to live with multiple variations of 'holocausts'.

ALEXANDER MICKÉ

Salmansdorfer Str. 94c,
A-1190 Vienna, Austria

SIR—One of your major arguments against the Vatican Instruction on *in vitro* fertilization warrants comment.

The point is made that "in laboratories, for example, it is important to know when the development of artificially fertilized embryos must be stopped for fear of accumulating unwanted and inviable human individuals in Petri dishes". You go on to suggest that it would also have been prudent on the part of the Church not to have taken an absolute stand in defending the inviolability of the life of the embryo from the moment of conception, as experimental evidence may some day be forthcoming that will allow us to say with certainty at what point in its development it becomes a new human individual.

The argument applies even more strongly in reverse. If one is in doubt as to when one can say that a new human individual is present, and if one holds "that the lives of individual human beings should be respected absolutely", the prudent course is to treat the embryo with absolute respect from the moment of conception.

An analogy, perhaps somewhat crude, with the question of when an embryo should be accorded the respect due to a human being may be drawn with an aspect of my growing up in farm country in western Canada. When we were young, we used to go hunting with rifles or shotguns. But, in addition to teaching us the technicalities of how to use the guns, our parents thoroughly inculcated us with an elementary lesson of prudence to avoid accidentally shooting another person. It

was to the effect that we must never shoot at a movement or a figure unless we were absolutely sure it was not a person. Lack of obedience to this simple principle was (and still is) responsible for many hunting accidents and deaths. (In this, I prescind from the debate over hunting for sport, which I have not done for over 30 years.)

This argument also applies to the opposition, from the Roman Catholic Church and others, to clinical abortion at any stage of pregnancy, namely that one cannot take the risk that one may be killing an innocent human being.

JOSEPH G. ATKINSON

Merck Frosst Canada Inc.,
CP/PO 1005, Pointe-Claire-Dorval,
Quebec H9R 4PE, Canada

What Fisher meant

SIR—The basis for thinking that R.A. Fisher would have approved the optimization methods underpinning the modern evolutionary approach to biology¹ is his statement² that: "The vital statistics of an organism in relation to its environment provide a means of determining a measure of the relative growth-rate of the population, which may be termed the Malthusian parameter of population increase . . . The Malthusian parameter will in general be different for each different genotype, and will measure the fitness to survive of each . . . Any net advantage gained by an organism will be conserved in the form of an increase in population, rather than in an increase in the average Malthusian parameter, which is kept by . . . adjustment always near to zero." Fisher must have meant 'gene' when he wrote 'genotype' because the proof³ depends on that. Thus inferior genes are removed from the population and those left maximize the Malthusian parameter.

This process is subject to constraints, as Fisher well knew when he wrote the often-quoted words² that "it would be instructive to know not only by what physiological mechanism a just apportionment is made between the nutriment devoted to the gonads and that devoted to the rest of the parental organism; but also what circumstances in the life-history and environment would render profitable the diversion of a greater or lesser share of the available resources towards reproduction". Thus did Fisher provide the twin principles, of fitness maximization, but subject to constraints, which underpin the optimization approach.

R.M. SIBLY

Department of Pure and Applied Zoology,
University of Reading,
Reading RG6 2AJ, UK

1. Edwards, A.W.F. *Nature* 326, 10 (1987).

2. Fisher, R.A. *The Genetical Theory of Natural Selection*, 2nd edn (Dover, New York, 1958).

3. Sibly, R.M. & Calow, P. *Physiological Ecology of Animals* (Blackwell, Oxford 1986).

The boycott of South Africa

Nobody approves of apartheid in South Africa, but continuing efforts by the rest of the world to ostracize South African science create as many problems as they may yet solve.

Johannesburg

STATE President P. W. Botha's speech on Tuesday last week, 19 May, has disappointed the sprinkling of optimists but confirmed the greater throng of pessimists in their disappointment. With his National Party re-elected as the majority in the White chamber of the tricameral parliament (there are separate chambers for Indians and 'Coloureds') on the slogan "Reform, Yes! Surrender, No!", Botha might have said something about his vision of reform, might even have delivered the much advertised "Rubicon" speech he failed to make just over a year ago.

In the event, Botha chose to make what the newspapers here have called a "low-key" speech, a reiteration of the familiar message that law and order are prerequisites of reform, the National Party's name for the dismantling of apartheid in some still unspecified way. On the following day, Wednesday, the Johannesburg police arrested more than a score of people in the city, some students included, saying that they were suspected of terrorist crimes.

Acknowledgements

ARRANGEMENTS for the journey on which the accompanying material was collected were made by the offices of the South African Council of Scientific and Industrial Research (CSIR) in London and Pretoria at the insistence of the author. The identifiable direct costs of the journey to and within South Africa were met by *Nature*.

The writer (John Maddox, editor of *Nature*) is grateful to the presidency and staff of CSIR for many frank conversations, as many splendid meals and for transport in and around Pretoria. He also wishes to acknowledge the assistance of the vice-chancellors of the Universities of Cape Town, Pretoria and the Witwatersrand; members of administrative and academic staffs there and at the Universities of Stellenbosch and Western Cape and at the University of South Africa; the Royal Society of South Africa; the editor of the *South African Journal of Science*, Dr Graham Baker; the directors of the Human Sciences Research Council, the Medical Research Council and the National Accelerator Laboratory and their staffs; and the directors of Altech SA and of the Anglo-American Corporation. □

What follows is a report of impressions (and facts) gathered during a ten-day scamper between Pretoria, Johannesburg, Cape Town and some of their environs between. It is offered in the hope of informing the continuing debate in the scientific community outside South Africa about the extent to which, and the manner in which, South African science is to be ostracized on account of the generally detested policy of apartheid.

Although this journal takes the line that the exclusion of South African scientists from international meetings such as last year's archaeological congress at Southampton is a kind of scandal (see *Nature* 319, 85; 1986), those who do not share that opinion may also have a passing interest in an account of how members of the South African community have reacted to recent events, not least to the 6 May election for members of the white community alone. In any case, ghoulish though it may seem to people in South Africa, there can hardly be a more intricate and absorbing problem than that which now confronts them.

The issue of whether there should be a boycott of South African science, and if so how, is also a conundrum for those who live elsewhere. Several mechanisms have been proposed, some of which are patchily in operation (see p. 272). A recent innovation is the suggestion that the boycott of South African science should be "selective", applicable to some but not to others. So which authority decides whether so-and-so is acceptable? And could the scientific community survive the precedent of its agreement that the right to contribute to the process of discovery should be determined on other than scientific grounds?

The purpose of a boycott (used as a portmanteau term referring to restraints in general) also needs closer definition outside South Africa. Several purposes are logically admissible. They range from making members of the South African community think more deeply about the social and political problems that confront them (and be more energetic in their resolution) to the effecting of political change in South Africa. But sheer distaste cannot qualify as a logical justification of boycott; if it did, prejudice would be licensed for the rest of time.

The simple safeguard is that boycotters should be clear in their own minds, and in advance, of the conditions in which they would relent. They also owe it not merely

to those affected but to their own objectives to declare what they are hoping to accomplish: abolition of the Group Areas Act, the abolition of the Population Classification Act, the attainment of an average non-white income exceeding that of whites, one person-one vote or something else? Not to specify the circumstances in which redemption would be possible must surely be tantamount to an unnatural punishment, in the language of the US Constitution.

The matter is important, and not merely for those who live in southern Africa. The scientific community there is an important part of the international endeavour; most of even those who would ostracize it now would probably regard its return to the fold as an unmitigated benefit. But failure in that regard would be synonymous with failure of the reform process (now apparently in stagnation) and much more serious trouble for all of us.

To the extent that the objective of boycott is to effect political change, it is inevitable that any account of its effectiveness must also seem political, perhaps tendentiously so. But what follows differs most from most contributions to this journal in being an attempt to reflect what people (mostly South Africans) think and to suggest what they (and the rest of us) might do (see p. 259).

There is a precedent. Seventeen years ago, accompanying a report of the condition of science in South Africa at the time, *Nature* likened excellent South African science to a Trojan Horse (*Nature* 228, 301; 1970). Surely rationality would triumph over obscurantism? And surely the growth of gross national product, bred of the application of science in technology, would undermine the integrity of apartheid? That hope has been mostly but not entirely disappointed; the science is even better, but the issue of apartheid is still there.

Much of the trouble is that the South African intellectual community, while more solidly and explicitly opposed to apartheid than previously, has not yet found a way of following the earlier prescription that it should seek a way "to argue the particular as well as the general case against the present arrangements" so as to "change the present climate". But if, 17 years ago, it seemed as if the greatest danger was that "time is running out", how can one by now be sure that it is not already too late? □

Legacy of history

Several degrees of separateness

APARTHEID means separateness in Afrikaans. As a political instrument, it was invented in South Africa in 1950, with the decision of the then government that people should be classified by race and then, if possible, kept separate from each other. This does not imply that whites and non-whites never work together; Afrikaans tradition is replete with the legend of the bond between the isolated farmer and his workers. It is merely, in the same tradition, that the two are racially and culturally distinct, and that their development will be harmed by too much mutual influence. "Separate development" used to be a common euphemism.

Apartheid is less obtrusive than it used to be, which does not imply that it is any less objectionable. But outwardly, there have been striking changes in 17 years. Physical segregation has disappeared during the working day. Your neighbour at a hotel bar may as well be non-white as white (but is likely to be white if only because of the persisting disparity of incomes). Racially mixed audiences will be mixed randomly, not segregated. And there has been a vast improvement of interracial manners; people shake hands with differently coloured people and introduce even their menials properly, as "Mister" so-and-so.

Less obviously but more significantly, some important elements of apartheid have been abolished. The requirement that non-whites should always carry an identity card (pass-book) has gone, together with those sections of the Immorality Act that, until three years ago, made interracial marriage a criminal offence.

There has also been a significant improvement of the economic condition of the non-white communities. One of the influences in that direction is the Sullivan Code (named after the black US religious minister who was denied a visa to revisit South Africa only two weeks ago) by which a large number of South African employers have bound themselves not to discriminate by race in the promotion of employees and to pay equally for equal work. (One company director says that the second provision, "an increase of costs but not of productivity", may have contributed to the recession of the past few years from which the South African economy is only now recovering.)

The increased flow of funds into the non-white communities has also had the effect of stimulating what Professor Johan G. Garbers, president of the Human Sciences Research Council, calls the "informal economy" of the townships in which the non-white urban labour force is segregated at night. One-man businesses are multiplying. The informal taxi-service

between the city and the surrounding townships, operated by the entrepreneur owners of mini-buses, is one tangible sign of that. Some successful entrepreneurs have even bought their way illegally into apartments in Johannesburg suburbs such as Hillbrow (from which the government, just a few days after the election, was threatening to evict them).

This is also the first year in which the number of black Africans sitting the school-leaving matriculation examination will exceed the number of whites. No doubt more of them would succeed if the dates set for this year's examinations did not include 16 June, known in the townships as 'Soweto Day' and usually marked by a non-white strike or 'stay-away'.

But if the educational provision for non-whites has improved, it remains a pale shadow of that for the children of the whites. Much the same is true of health care. But the provision of housing for non-whites has deteriorated under pressure from a population growing at 3 per cent a year; estimates of the shortage of houses for non-whites range from 700,000 (National Institute of Building Research) to twice as many.

Other features of persisting apartheid are more obvious. South Africans themselves appear not to appreciate that even the vocabulary of apartheid gives offence. Although people of all races may now sit together in bar-rooms and lecture rooms, the general and repeated use of nouns such as "white" and "coloured" (for a person of mixed descent) is a constant reminder of apartheid. Mercifully, "the bantu" (plural as well as singular) has been replaced by "black". These terms are used reluctantly in what follows.

The longstanding confusion about the status of oriental people persists; Japanese residents of South Africa have full civil rights, but although Chinese are regarded as 'honorary whites' and are sometimes able to vote in white elections, the right is denied on other occasions. The general usage of these terms is a reminder of the longevity of the Population Registration Act of 1950 and a standing invitation to make over-simple generalizations about the properties of social groups.

But in the cities, the most obtrusive features of apartheid are the manifestations of the Group Areas Act, which requires that non-whites should not live in the areas set aside for whites, but rather in separate locations which are often, in the Transvaal, many miles from a person's place of work. Even without conversation, the sight of the industrial suburbs at the end of the working day gives the show away. The sidewalks are crowded with non-whites, even those who work in

Sullivan Code factories, trekking to the nearest railway station, which itself may be a mile or so away, so as to travel home.

It needs little imagination to appreciate the indignity of the general rule that there is one form of locomotion for the whites and another for the non-whites. It is harder to appreciate the sense of indignity that, for the majority of urban blacks, comes from knowing that home is not really home at all, but only a temporary lodging away from the real home, which is a tribal homeland somewhere.

Conversation yields more pungent comments. A coloured worker in Cape Town makes what seems to be a common complaint, not wryly but angrily, when he asks why, if the donor of the heart used in the first human transplant operation at the Groote Schuur Hospital was non-white, do "they" insist on segregated cemeteries even when we're dead?

To be fair, the government seems to be recognizing that the homelands policy is collapsing under the weight of the social problems it has created. After a virtual halt to housebuilding in the townships in the 1970s, non-white urban settlements are growing again. Contracts are being let to civil engineering companies to install concrete foundations, water supplies and rudimentary sewerage, on up to 5,000 building plots at a time, together with the associated service roads and infrastructure. The idea is that coloureds and even blacks (who have been flocking to the Cape) should be able to build the houses of their choice on the concrete slabs. But the decision to accommodate more non-whites in Cape Province is not synonymous with the abandonment of the homelands policy.

In the regulation of where people live, the government may even become tougher. The Group Areas Act segregates living space, but the police act on infringements only on complaints by neighbours and, hiding behind their reputation for bumbledom, may choose not to act quickly. Over the past few years, the infiltration of blacks into two Johannesburg suburbs without retaliation has been taken as a sign that the authorities were prepared to let the act lapse, at least in some places. That could be another disappointment.

Startled to find that the Conservative Party, which will be asking that the government should enforce the existing apartheid laws, had won 25 of the votes in the general election on 6 May, the government was leaking news of an impending 'crack-down' by the following weekend. Those found by the courts to be in violation may be evicted (though a court decision means eviction is allowable only when there is alternative accommodation in a legal area). The ultimate penalty for transgression is that illegally occupied property may be confiscated by the authorities and sold; the owner does not

Isolation forces the pace for cyclotron

Now, after 22 years of haggling, South Africa has an accelerator of its own. This modern if slightly passé symbol of nationhood occupies more than 1,000 hectares of the flat sandy Pleistocene infill of the Western Cape, within sight of Table Mountain (on a clear day) and midway between two of the chief user groups at the Universities of Cape Town and Stellenbosch.

The machine, a variable-energy cyclotron capable of producing 200 MeV protons now and comparable beams of heavy nuclei when the necessary ion sources are ready, is home-designed and home-built, although some components have been bought in from abroad. Director D. Reitmann and his sponsors, the CSIR, rightly boast that the cyclotron produced a proton beam when it was first switched on three months ago, but the former will not disclose the cost (R135 million so far), saying that the effect would be to "start all those arguments that the money should have been spent on schools or housing".

The long saga of the cyclotron typifies both the difficulties and the rewards of self-sufficiency engendered originally by pride and, latterly, by isolation. Eleven years, from 1966 to 1977, is a long time for a feasibility study of a machine that is an elaboration, even if a good one, of a well-known principle. The cost escalation, from R35 million over five years to four times as much over twice as long, is partly but not wholly attributable to inflation, financial crises and the usual inconstancy of governments.

The rewards are also plain. Reitmann points with pride to the home-made sections of the electronic racks that control the current in the coils that shape the magnetic field, as a function of radial distance, in the four sectoral magnets. The machine will be used by both nuclear physicists and physicians interested in

necessarily receive the proceeds.

So when will there be 'one man, one vote'? Significantly, none of the white parties fighting for election on 6 May appears to have raised this as an issue that must at some time be decided. To have done so would have been a recipe for losing votes. It seems to be widely agreed that the Progressive Party lost ground (so that it has yielded place as the largest parliamentary opposition to the Conservative Party) for hazarding the less radical suggestion that it would negotiate with the banned African National Congress (ANC), which organizes anti-apartheid activity, including bombings, from its base at Lusaka, the capital of Zambia.

So great are the electoral dangers of this forbidden question that it is unlikely to be mentioned even in the next election, due



The cyclotron in the National Accelerator Centre — no route to international respectability.

particle beams for cancer therapy. There is also a facility for the production of short-lived isotopes. Reitman says that his cyclotron is matched only at a few laboratories around the world.

Who will use the machine? There are strong nuclear physics groups at nearby Cape Town and Stellenbosch and also at the University of the Witwatersrand. Time will be allocated by a users' committee, but the laboratory will provide all but the most experiment-specific ancillary equipment. Reitmann says that he would also welcome physicists from overseas "to repay the hospitality we've received all these years". He says there have been some enquiries from US investigators, and suspects there would have been more if his colleagues had not been prevented from singing the praises of their machine at last December's cyclotron conference in Japan, from which they were banned.

two years from now (but which could be postponed if the government instead embarks on another round of constitutional change, the creation of a black council of state, for example).

Even liberal South African critics of apartheid do not seem to regard universal franchise as a necessary component of what the optimists call the "post-apartheid society". The white population, already a minority, knows that it will be an even smaller proportion of the total as the years go by. So people talk instead of elaborations of the tricameral parliament now in force, or the more ingenious electoral system devised by last year's *indaba* in Natal, which would ensure that no community's constitutional rights could be usurped.

One person, one vote, tomorrow would

The medical case is less convincing. Some at a nearby medical school wonder whether the cost of the hospital will be justified by the treatment of just over 100 patients a year. A more practical difficulty is that nuclear physicists' interests may soon outgrow this machine, when an understandable feeling that the hospital should be decently amortized will make it a more durable incubus on the budget of the research foundation FRD, to which it will be transferred on 1 April next year.

All that could be different if, five years from now, South Africa had rejoined the international community, when international swapping of machine time might make it possible to sustain a diversified programme of academic research. As it looks now, Reitmann's splendid cyclotron will not be the price of entry to international collaboration but a monument to ambitions now frustrated. □

indeed be a recipe for disaster. The non-white majority has no political structure within which to organize its opinion. Moreover, with the present segregation of people's home addresses, it is unthinkable that the outcome of a free election could be determined on other than racial lines. The pity, at present, is that there is probably still a proportion (but shrinking) of non-whites willing to take part in the white political structure. A poll by *The Sowetan* in advance of the all-white election suggested that even Mr P.W. Botha's party could have counted on some support in Johannesburg's largest black township. The obvious hazard in the way of peaceful change is that non-white opinion will have hardened against participation in the post-apartheid society before that comes into being. □

Scientific publications

Boycotts are biting hard

"WHAT hurts us most is that we feel we've been let down by our friends." Thus a senior official of CSIR last week, brooding about the way in which his colleagues run into difficulties when seeking to attend conferences overseas, when they try to organize meetings in South Africa and when they try to decide in their own minds whether a paper submitted to an international journal has been rejected on scientific grounds or for a more sinister reason.

Last year's disinvitation of South African scientists previously expected at the Southampton Archaeological Congress became a controversial issue because the exclusion of at least one of the disinvited, Dr Phillip Tobias, the palaeoanthropologist who is professor of anatomy at the University of the Witwatersrand, was plainly a disservice to the other members of the congress.

Since then, South Africans have found that it has become even more difficult to attend conferences overseas. Some of the problems are administrative. Japan, for example, has become virtually out of bounds, issuing visas to South Africans travelling for business purposes, but not to tourists; conferences appear in the eyes of consulates to fall vaguely in between. Both Canada and Australia now require that South Africans seeking a visa should present themselves in person at their consulate in London (or in Harare, the capital of Zimbabwe, in the case of Australia). Applicants are advised to allow three days, at the least adding to the length of any journey. (One of those affected, eventually given a Canadian visa in 20 minutes, said that he had nevertheless found the three days spent in a London hotel an opportunity to prepare his paper carefully.)

Administrative hassle is especially onerous for South Africans working for government research establishments and wishing to visit the United States. Organizations such as CSIR are classified as "para-statal". Those who work for them who apply for a US visa are required to specify in advance the names and addresses of those whom they plan to visit, giving a full advance account of where they will be at all times during their visit to the United States. Some have been embarrassed that their hosts have been telephoned by the Federal Bureau of Investigation, checking that the declared schedule has indeed been followed.

For the rest, people's participation as individuals in conferences to which they have been invited yields a patchwork of experience. Sometimes people are allowed to attend but are asked not to present papers, on the grounds that to do

so would invite protests. Sometimes they are invited and then asked not to attend. Holding conferences in South Africa is much more difficult. Among conference organizers, it is a common experience that people from overseas agree, at first, to attend, but then change their minds, often at the last minute, giving as explanation pressure from colleagues or fears for their personal safety. A medical conference at Cape Town last year was ruined when eleven out of twelve overseas speakers failed to appear. Nuclear physicist Professor F. Sellschop, a vice-principal of the University of the Witwatersrand, says he "has nightmares" on account of his obligation to organize a conference in South Africa next year.

In many ways, the increasing difficulty of holding international conferences in South Africa is the more serious impediment to academic life. Attending a conference overseas is a feather in one's cap (and an entry in one's *curriculum vitae*), but visiting speakers at conferences in South Africa are often the only means by which younger scientists can be given a sense of swimming in a bigger pool.

The supply of journals is another growing problem, chiefly because some Scandinavian publishers are no longer allowed to supply journals to South Africa. This, for example, applies to *Acta Crystallographica*, the journal of the International Union of Crystallography, itself affiliated



Dr C.F. Garbers, CSIR President, has recruiting problems.

to the International Council of Scientific Unions whose committee on the free circulation of scientists has traditionally taken a strong line on restrictions of this kind. (The publisher is the Danish company Munksgaard.) The practical effect in South Africa of restrictions on the supply of journals has been to increase library costs and further to delay their arrival. For a community already isolated by geography, the results could be serious. "Without our journals, we are dead", says one academic.

Restrictions on the more permanent movement of people are more serious, but

growing. CSIR is one of the organizations chiefly affected. The president, Dr C.F. Garbers, says that the organization used to recruit the equivalent of "one institute a year" from overseas, roughly 100 people of all grades, but that this has now shrunk to two or three individuals a year. But even the English-speaking universities with reputations for liberality are going through a bad patch; overseas recruitment at the University of the Witwatersrand is down to a third of normal.

Although political uncertainties are important factors in the attenuation of overseas applications for posts in South Africa, it is probable that the recently depressed value of the rand has played a part. "People turn the salary we offer into their own currency, not realizing that the rand goes a long way in South Africa", says one CSIR official. But the value of the rand matters to immigrants with overseas commitments; one CSIR researcher, with a salary of R75,000 a year but with US alimony payments to make, is on the point of returning whence he came.

The government organizations could be badly hurt if the trickle of departing researchers becomes a stampede. No fewer than 30 per cent of CSIR's scientists hold foreign passports. The corresponding figure for Mintek, the smaller minerals research organization, is more than 60 per cent. Given the difficulty of bridging the gaps that might be created by an exodus of foreigners from the output of South African universities, both organizations are exploring the possibility of an agreement with Taiwan (otherwise known as the Republic of China) that would allow for the secondment of Taiwanese researchers to South Africa on short-term contracts; an ironic snag is the persisting ambiguity about the status of people of Chinese origin in South African society.

This bears on a serious threat to South Africa's ambitions. By the scale of the white population (6 million) South Africa is a science-intensive economy, and has been able to sustain that state of affairs by overseas recruitment. But now there are worrying signs that fewer technically trained people are emerging from the white population as young people are tempted to more lucrative work in other fields and as the demographic pool of young people shrinks; the number of engineers graduating from South African universities, now roughly 1,000 a year, will fall to 700 in 1990 for example.

There are also signs that the inclination of new graduates to find first jobs overseas, not to return, is strengthening, while the output from the universities of non-white technically trained people falls far short of the need. (Organizations of all kinds are now eager to recruit technical graduates from the non-white population, but CSIR has at present only a handful of such people, while the director of one

CSIR institute says that there is an element of "tokenism" in the process.) Anxiety about the future supply of technically trained manpower is a sufficient — although not the only — explanation for the conviction in the technical community that non-white education should be one of the government's priorities.

Inevitably, these growing restrictions have increased the sense of isolation from which South Africa suffered, for geographical reasons, even without restrictions on the movement of people. But the issue that most offends those working at South African laboratories and universities is the belief that they will soon be denied access to overseas journals.

Wily, people are ready generally to acknowledge the temptation to ascribe the rejection of articles submitted to overseas journals to prejudice by editors; easier that than to acknowledge that the article may not have been up to scratch. But there are some authenticated cases of articles that have been rejected merely on the grounds that they had been submitted from South Africa.

The most spectacular such case so far to come to light is the rejection of an article submitted to the *Zoological Journal of the*

play an important part in the society's affairs. No further explanation was forthcoming by the end of last week.

Other authenticated cases of papers rejected on the same grounds come from Scandinavia. Thus the Norwegian journal *Holarctic Ecology* in February 1986 rejected, "against the wish of the editors", a contribution from Dr Richard S. Knight of the University of Cape Town on the grounds that "due to governmental decisions", there is a ban on the publication of South African articles "in journals sponsored by governmental institutions (in our case the Natural Science Research Council)".

More confusingly, the ornithology journal *Ornis Scandinavica* explained to Dr D.C. Duffy (also from Cape Town) that "we cannot handle contributions from South Africa", but asked if Duffy might not use his North American address because "the editor has now read your MS and found it interesting".

Although the policies of the Scandinavian journals have not, hitherto, been much remarked upon, proposals for a boycott of South African research publications are not new. Thus, in 1985, a group of researchers at the University of Wisconsin urged that the Ecological Society of America should not in future publish articles dealing with research carried out in South Africa, whether by South Africans or others, on the grounds that doing so would confer legitimacy on apartheid (see Daniel W. Schneider *et al.*, *Bull. Ecol. Soc. Am.* 67, 161; 1986 and references therein).

As seen from South Africa, journals differ markedly in their policies, although those of the Scandinavian journals (which include a number of important environmental and ecological journals) are the most uniform. One particular concern is that the boycott might spread to the more important Dutch publishers. The Netherlands has been so zealous in the enforcement of restrictions that Professor H.W. Rossouw, vice-rector for academic development at the University of Stellenbosch, a philosopher who took his PhD at the Free University of Amsterdam, now says that while he is able to visit his Dutch friends at their homes, he cannot be invited by them onto university premises.

The basis of these restrictions is the resolution of a meeting of the ministerial meeting on political cooperation of the European Communities on 10 September 1985, which decided that member states would "harmonize their attitudes" on "discouraging cultural and scientific agreements except where these contribute towards the ending of apartheid or have no possible role in supporting it...". This declaration is essentially the same as that of the Nassau meeting of the (British) Commonwealth Heads of Government in October 1985.

These declarations probably also explain why continuing collaboration between a British scientist at the Harwell laboratory of the UK Atomic Energy Authority and a colleague in South Africa is now conducted by private mail, and why the British Foreign and Commonwealth Office last year declined to allow Dr Nigel Bollard, of the British Antarctic Survey, to act as a referee of a South African examination of the environmental effects of the now-abandoned project to build an airstrip on Marion Island, part-way to Antarctica.

The South African reaction to developments, apart from a sense of hurt, may be summarized as follows:

Boycotts are bad for science. Most people would probably agree.

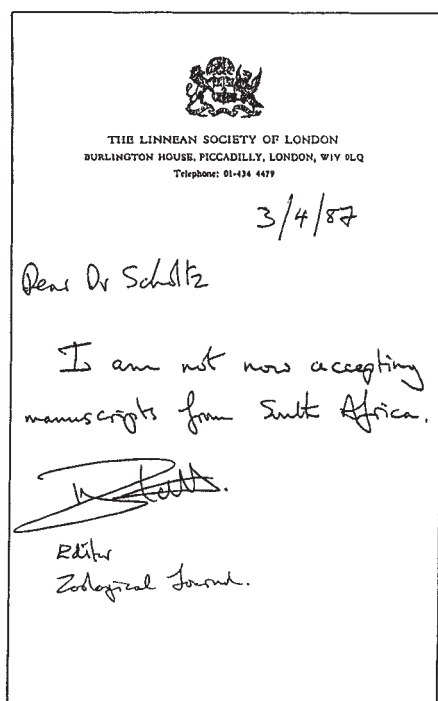
Damaging South African science now will impair the capacity of South African science to contribute towards the solution of the problems of Africa north of the Limpopo. The flaw in this argument is that there is no immediate prospect that African states other than Botswana, Lesotho and Malawi will readily accept the technical assistance South Africa is eager to offer, although there is a formal agreement on housing development with the Islamic Republic of Comores (the island state north of Madagascar). There is also a small trade in technical documents between research organizations in South Africa and African states to the north, using mutual friends in neutral countries such as Britain as post offices.

Damaging South African science now will impair its contribution to the 'post-apartheid society'. The snide answer to this is that the point would be more telling if it were more evident that there will ever be a post-apartheid society. More seriously, this is a line of argument that will weigh differently with different people.

Severing ties will diminish external influence on our affairs. Unfortunately, this argument cuts both ways. It seems also to be agreed that the effect of the boycott so far has been to make the South African community think harder about domestic policies.

One problem is that it is hard to tell where boycotts lead. Some directions are suggested by the case of a Cape Town biomedical scientist collaborating on a grant application to the US National Institutes of Health who felt bound to agree to his US collaborator's suggestion that success would be more likely if his name were suppressed.

On the other side of the coin is the case of entomologist Professor Jan Giliomee from the University of Stellenbosch, who says that he was able to shake off the influence of his conservative Afrikaans background only by spending two years in London "seeing that people could mix together without harm". □



Curt rejection — so far unexplained.

Linnean Society by Professor C.H. Scholtz of the Department of Entomology of the University of Pretoria. The curt rejection letter from the editor (see accompanying illustration) says "I am not now accepting manuscripts from South Africa".

The circumstances are odd. A spokesman for the Linnean Society said in London this week that the council had determined that it would not discriminate against publications from South Africa, adding that at least two South Africans

Three kinds of universities

Source of unrest and anachronism

WHATEVER the delay in ridding South Africa as a whole of apartheid, the university system is moving steadily in that direction. Two universities, Cape Town and Witwatersrand (in downtown Johannesburg), have always sought to be open to students and scholars of all races and persuasions. Others, notably the Afrikaans Universities of Stellenbosch and Pretoria and the Rand Afrikaans University, are now following suit.

Witwatersrand and Cape Town vie for the reputation of being in the van, educationally and even politically. Certainly they both now enjoy the doubtful distinction of being embattled campuses. At each of them, in the days preceeding the election of 6 May, a student demonstration provoked police intervention. Professor J. Reed, deputy vice-chancellor, is proud of having seen Cape Town students refrain from physical violence as a police truck moved slowly along their ranks taking television pictures. At Witwatersrand, Dr Karl Tober, the Austrian-born vice-chancellor, was able personally to persuade the students to make their gathering legal by holding it inside the university.

University education for non-white students is formally regulated by the 1959 Extension of University Education Act, under the terms of which separate universities were established for the four principal racial groups in South Africa (white, black, coloured and Indian). The same act gave the government power to restrict by quota the proportions of other kinds of students attending particular universities, a system that remained in effect until 1983 (and, being still part of the country's administrative law, could in principle be reimposed at any time).

The quota system, although not now enforced, seems to both universities to be an infringement of their freedom. Both have taken advantage of the respite from restriction to increase the proportions of non-white students to roughly 18 per cent (out of 13,000 students at Cape Town and 18,000 at Witwatersrand). The result is evident at any public place at which students are gathered, drinking coffee or whatever: apartheid might never have been. But there is a price to pay. Educationally, given the inadequacies of non-white secondary education, students are not admitted solely on the basis of their performance in the national school-leaving examination. Instead, and to the chagrin at least of the white students who fail to gain admission to these institutions, admission is determined both by performance and an assessment of potential. Witwatersrand was the first to institute a student support programme to help non-white students with the problems of learn-

ing to learn. Those concerned usually spend two years on the content of what would normally be the first of three years of study.

A further cost of open access (for academics as well as students) is that both universities have been driven to violation of the law, and of the Group Areas Act in particular. Tober argues eloquently for example that "sending black students back to Soweto every night" would guarantee their academic failure. The result is that black students are encouraged to live in university residences, which are racially mixed on the grounds that that is both educative and a further educational support for non-white students. One of the most conspicuous violations is represented by the student dormitory in Johannesburg built with a grant from the Anglo-American Corporation, the gold-mining company. Cape Town's violation is that non-white academics are being allowed to purchase by means of mortgages houses in whites-only areas owned by the university.

Another price that has had to be paid may be the radicalization of the student-body (and even, perhaps, of the university itself). Non-white students, to the regret of their teachers, have chosen to form a separate representative organization, with which the vice-chancellor seems perpetually in negotiation; two weeks ago, for example, he found it necessary to spend the best part of the morning trying to dissuade the non-white students from lending their support to an industrial dispute between the university's manual workers and the administration. Tober says the effort is exhausting.

The drumbeat of criticism from outside is also debilitating. Making no secret of its opposition to apartheid, Witwatersrand is repeatedly criticized as a hotbed of radicalism by politicians, in letters to the newspapers and, sometimes, even by the newspapers themselves. Ironically, the latest row concerns the university's decision not to allow Mrs Helen Suzman, perhaps the best-known member of the liberal Progressive Party, to speak on the campus in advance of the 6 May election: Tober says that it is not the university's policy to allow politicians to use the campus for political purposes, members of the Progressive Party blame him for having helped the party to its electoral defeat while others accuse him of caving in to a demand that the university should boycott the election altogether.

Radicalism is also a problem at Cape Town, where the latest issue is whether a new wing of the medical school's teaching hospital should be advertised publicly as a racially undifferentiated institution. The

facts are that the existing hospital is organized in such a way, while the design of the new wing is such that segregated treatment would not be feasible. So why not say so, loud and clear, "to let people know where we stand"? The obvious rejoinder is that Cape Town is in no need of further trouble. But there are academics as well as students who will not let the issue lie.

The Afrikaans universities are calmer places, but not necessarily less committed to reform. Stellenbosch, for example, by tradition a bastion of conservatism, has recently made a name for itself as an outspoken advocate of change, not so much because the university has taken a position on the subject, but because a group of academics calling itself "Discussion Group '85" has taken to making statements on the urgency of the present situation in South Africa. The group says it has no connection with political parties, but the coincidence that Dr Denis Worrall, lately South Africa's ambassador in London, should narrowly (by 39 votes) have failed to beat a potential successor to P.W. Botha for a neighbouring parliamentary seat is thought more than mere coincidence.

Like the others, Stellenbosch has begun taking in non-white students, beginning with its graduate schools. Now there are a few hundreds (out of a total of 13,000), but the university's immediate objective appears to be to strengthen its graduate schools and its research, each of which may help to balance the impending decline in the numbers of white school-leavers.

There has been much progress. Stellenbosch has quickly become strong in electronics, in part with help from contracts with the military. But Professor John H. Cloete, of the Department of Electrical and Electronic Engineering, a reformer and a critic of the scientific boycott, asks why the outside world expects South Africa to be willing not to defend itself.

Stellenbosch has also been surprisingly ready to set up novel academic groups — there is an institute of biotechnology (with help from the Anglo-American Corporation) and centres for economic analysis and hard-headed (which means economic) future studies that have contributed much to recent introspection about the future of South Africa.

Professor H. W. Rossouw, vice-rector for academic development, is also keen to emphasize traditional strengths — agriculture is strong while Stellenbosch has the only department of forestry in South Africa. He claims that Stellenbosch is helping South Africa to solve within one frontier a problem that also confronts the world as a whole, that of reconciling the needs of an advanced (white) society and one that belongs to the 'third world' (and which is black).

Stellenbosch in temperament is much

more like a university in the United States or Western Europe than are Cape Town and Witwatersrand. Theoretical physicist Professor C. A. Engelbrecht, one of the instigators of Discussion Group '85, argues for the importance of accommodating within a university scholars with opinions at odds with the majority, implying that the English-language universities have inherited Stellenbosch's old reputation for conformity (but to a different cause).

Unlike Cape Town and Witwatersrand, Stellenbosch is not in the business of affirmative action on behalf of non-white students, requiring that they should be qualified for admission and willing to be taught in Afrikaans (which usually means that they are coloured, not black). But how can a mere half-dozen Afrikaans universities sustain a viable academic community across all disciplines? Rossouw acknowledges that "we know we have to become dependent on, and a part of, the English-speaking world".

The University of Pretoria, the largest (with 21,500 students) in South Africa except for the University of South Africa, with 92,000 students, mostly on correspondence courses, is in the same mould. Knowing that the government has promised to leave graduate education alone, Professor D. M. Joubert, the rector, hopes to increase the proportion of graduate students from a quarter to more than a third. The number of non-white students has grown steadily (to some hundreds) since the university became open, but with the reservation that entrants should be properly qualified and willing to be taught in Afrikaans, but there is also now a plan to start a student support programme (not exclusively for non-whites).

Again there is a certain uneasiness about the extreme positions being adopted by the English-language universities, although Joubert, now chairman of the committee of university principals, says that the South African university system "has more cohesion than ever". To those who ask why the pace of social change is so slow, he says that South Africa's equivalent of the Deep South in the United States is the "Far North".

The University of the Western Cape, by contrast, is in temperament more radical even than the English-language universities. The University of the Western Cape (just under 10,000 students, mostly coloured) lies at the other end of the spectrum. Professor Chris Johnson, head of the Botany Department, asks that people should understand the difficulties of a university (and of its academics) which, by purpose, exists to serve disadvantaged students. Nevertheless, he says, the university has ceased to be the "glorified high-school" it was at the beginning, and after a quarter of a century now recognizes that students' learning must be coupled

with political awareness if they are to play a part in the future South Africa.

Western Cape happens to be in favour of an academic boycott of South Africa, but selectively. At the end of last month, the university senate voted not in future to invite to the campus people who did not "show solidarity with its commitment to the struggle for a non-racial democracy in South Africa". Acknowledging that the decision might seem to compromise academic freedom, the senate said that refusing to invite certain foreign academics would demonstrate "resistance against present structures". Two weeks ago, after a month-long strike by dental students culminating in a campus-wide strike, and after the report of a committee of inquiry, the university also decided to terminate its appointment of Professor

Jeffrey Cohen, who had been accused of racism.

Those elsewhere who believe they remember the language from the student movements of the 1960s in Western Europe and North America may be tempted to write off these developments as passing phenomena, but that could be mistaken. One difficulty is that the students are apparently united in their opinions (and that there are a lot of them). Another is that they have the support of a substantial body of their teachers. Yet another is that Western Cape was meant to be the pinnacle of education for coloured people in a society built around the doctrine of separate development. The architects of that policy must wonder whether another doctrine might have yielded a more comfortable result. □

Research council

Facing self-imposed turmoil

As if external troubles were not enough, South Africa's major research organization, CSIR (for Council of Scientific and Industrial Research), is in the thick of a radical upheaval. The plan is to live within a subsidy from the government fixed for the foreseeable future at, what it is at present, 234 million rand. With inflation at 16 per cent a year, radical steps are plainly needed.

CSIR is not alone. The government's other statutory research organizations, the Medical Research Council (MRC), that responsible for minerals research (MINTEK), human sciences and the South African Bureau of Standards were put in the same boat by a decision of the Minister of National Education last year. But the upheaval at CSIR, the biggest, is inevitably the most conspicuous. Moreover, CSIR has jumped at the chance to change with apparent relish.

The euphemism for the new state of affairs is "framework autonomy". The object is to make the research organizations less dependent on public funds, ensuring that they "enhance their orientation towards local (meaning national) needs"; in return, the organizations will be more free to spend their income as they choose. Flexibility on salaries is one goal.

How is it possible radically to change the direction of an organization employing some 2,000 researchers, already organized into a system of institutes? CSIR is probably lucky that most of its people are on the same site at Pretoria; they can the more easily be

moved into other groupings. And what is to happen to the organization's crucial function, over more than 40 years, of providing research grants for university researchers?

The intention is that academic research support should be hived off to a separate organization, called the Foundation for Research Development, for the time being within the parent organization, but CSIR has also said that it will transfer this whole function to some other organization if that is what its customers prefer.

By the regular repetition of the remarkable exercise carried out in 1985, when



CSIR deputy presidents Arndt (left) and Clark (right) are forced to look to industry for support.

potential grant recipients were qualitatively evaluated as researchers (and when the universities of Cape Town and the Witwatersrand were found to have 60 per cent of the outstanding people and more than a third of all other qualified researchers), the managers of this programme (run by deputy-president R. R. Arndt) hope deliberately to foster co-operation in research.

CSIR also manages facilities intended as national centres, the National Accelerator Laboratory, for example, but also

the Sutherland Observatory which has become one of the major observatories in the Southern Hemisphere. Since the decision of the British Science and Engineering Research Council to withdraw from the observatory, ostensibly on the grounds of economy (it used to pay 30 per cent of the running costs at Sutherland in return for 40 per cent of the observing time), South Africa has picked up the tab.

The reorganization appears largely to be the conception of Dr Brian Clarke, until two years ago director of the materials research institute but now a vice-director of CSIR. The programme for the months ahead is disconcertingly rapid, and perhaps meant as such. Almost a score of applied research institutes are to be rearranged into eleven, each of which has "technology" as the dominant noun.

The posts of director have been advertised publicly, and present incumbents invited to apply for posts in the new struc-

ture. CSIR's decisions will be known in mid-June, whereafter the new directors will spend two months training for their new responsibilities. There will follow a two-month period of more detailed assessment of what the new structure might expect to do for (and earn from) industry. Employees in the present structure will be told their fate by mid-November. The new arrangements will come into effect on 1 April next.

Reducing staff numbers is no part of the plan, but there are many at all levels who wonder whether there will be a place for them in the new organization. Dr E.N. van Deventer says that CSIR will be able to offer generous severance terms to those for whom there is no place. On the question whether CSIR is sure that it will be able to avoid the mistakes of other government research organizations seeking to bend research to industrial needs, he says simply "We don't know". □

of electoral colleges so as to prevent one community establishing dominance over the others, or the rights of locally elected authorities to relax the apartheid rules as they decide, appear to be the favourites, with the abolition of the Group Areas Act a second choice.

The most chilling opinions are those of a handful of non-white academics (echoed independently by one white researcher) to the effect that change will come about only if there is a discontinuous sharing of political power, most probably brought about by confrontation. Some put the point more in sorrow than in anger, regretting opportunities now lost. But others are very angry.

These views are not necessarily representative of non-white opinion generally, but in the absence of a mechanism for enabling all non-white opinion to be democratically expressed, nobody can know how far the non-white community is already radicalized.

One curious feature of what seems to be the more general opinion that reform is necessary, but that the recipe for reform is hidden, is that the intellectual community itself has been silent about the directions in which peaceful change might be engineered. This is why the Stellenbosch group has had such an influence.

In its statement on 7 March, the group asked for the abolition of the Group Areas Act, the statutory definition of race by the Population Registration Act, the tri-cameral parliamentary system and the Separate Amenities Act (now as much honoured in the breach as otherwise). The group also asked for an unambiguous declaration that political power would be shared with non-whites in a manner to be determined by negotiation.

It is also worth remarking that there has been a wealth of studies of the social problems of South Africa, by academic think-tanks, by public commissions and by private industry, typified respectively by documents put out by the Institute of Strategic Studies at the University of Pretoria, the Langer report on South African education which in 1981 put the case for a concentration of effort on non-white education (conducted under the wing of the Human Sciences Research Council) and the recently published report on pathways to a peaceful transition by Dr Clem Hunter of the Anglo-American Corporation.

So why has the academic community been so silent on a subject on which its future may depend? One explanation is that people have not been silent, but have spoken out in muffled code. Another, offered by one research administrator, is that "We South Africans belong to secret societies, not big ones like the *Broederbond* but small groups. My friends and I know what we think, but we do not like to tell other people." □

Intellectual opinion

What people are saying

HERE is a remarkable discovery — intellectual opinion in South Africa is despondent at the result of the election on 6 May. Most people are also impatient that the government hangs fire on reform, but there are some who expect nothing else and others who may welcome a state of affairs in which the present system is likely to collapse under the weight of its inconsistencies.

These are the impressions left by a series of remarkably open conversations with researchers and academics, individually and in groups, as well as with a small number of businessmen. They do not have the validity of formal opinion surveys, but that may be more than compensated for by the passion with which people frequently expressed themselves.

One sharp contrast between South Africa now and what it was in the 1970s is that there seems to be no great danger of running into eager declarations, by supporters of the present system, of the reasons why the pace of enlightenment must be slow. Only one (encountered in London) has offered the opinion that black people would remain disadvantaged for a long time "because there were living in mud huts one or two generations ago". Most academics holding this opinion tend to hold their tongues.

The most common view is that, because the National Party which forms the government has been startled by the strength of the vote for the Conservative Party to its right, and particularly by the fact that no fewer than 53 of its parliamentary seats are now held by majorities of fewer than 1,000 votes, it will be forced to be more accommodating to the conservatives and less attentive to reform.

Some white voters blame external pressure for this state of affairs, saying that South Africans are stubborn folk ready to retreat into the *laager* when they are attacked. Others blame "the West" for failing to give the government credit for the reforms brought about since 1980 and the general improvement of living standards. "Why is it always the stick and not the carrot?" Asked why the Common wealth's group of three "eminent persons" had been treated disdainfully in South Africa more than a year ago, one research administrator volunteered that "our government is often terribly tactless". Many people say they are convinced that the decision of the US Congress to impose economic sanctions reflects the attempt of the Democratic Party to capture the "black vote" in the United States, and that it will be a bad business if the Democrats win in 1989.

Some blame the frailty of other white voters, alarmed either by the security threat or by the Progressive Party's statement that it would negotiate with the African National Congress, implying a rapid transfer of power to the majority. Only one seeks seriously to see the result of the election in a cheerful light, saying that it is not the case that the electorate has moved to the right (against reform), but that the government party has moved in the opposite direction. That is balanced by a whispered "I'm terrified of what will happen".

Most white voters eager for reform are nevertheless unsure of what reforms they would like the government to attempt. Constitutional changes of the kind suggested last year by the Natal *indaba*, involving universal suffrage within a system

Terrestrial neutrinos reappear

Careful experiments in earthbound laboratories may be less exciting than supernovae, but will in the long run be more useful

THE 20 neutrinos that arrived on Earth from the recent supernova have prompted strenuous attempts by physicists and astronomers to extract the microscopic properties of the neutrino from the meagre data. This excitement may have caused some people to forget that laboratory experiments with similar aims have been going on for some time. As a reminder of these long standing efforts, the current issue of *Physical Review Letters* contains two somewhat contradictory contributions to the debate on the neutrino mass.

Physicists at the Institute of Theoretical and Experimental Physics in Moscow (S. Boris *et al.* *Phys. Rev. Lett.* **58**, 2019; 1987) describe new results which support their claim of 1981 for an electron neutrino mass of 30 eV, but another group (J.F. Wilkerson *et al.* *Phys. Rev. Lett.* **58**, 2023; 1987), working at Los Alamos National Laboratory, has data that set an upper limit of 27 eV. Experimental errors no doubt allow these two results to be reconciled, but the Los Alamos group hopes soon to increase the sensitivity of its method down to 10 eV, and the technical differences between the two essentially similar experiments are worth reviewing.

The basic idea is appealingly simple. Tritium has a beta decay whose products are a helium nucleus, an electron and an antineutrino, with a total energy of 18.6 keV. To do the experiment, collect some tritium, and record the energies of the electrons that emerge as it decays. Then plot a suitable graph of the number of electrons detected as a function of their energy. If the neutrino is massless, this graph is a straight line intersecting the axis at 18.6 keV, but if the neutrino is massive, the graph turns over and goes to zero at an energy of 18.6 keV less the neutrino's mass. Measuring the endpoint of tritium beta decay thus measures directly the mass of the neutrino.

Two problems arise in putting this idea into practice. Because the neutrino mass is at most a few tens of electron volts, a small fraction of the energy released in tritium decay, any departure from a straight line occurs very close to the endpoint, so good energy resolution and statistics are needed to reveal a turnover in the graph. (Tritium is used as it has the lowest energy release of any beta decay.) The ingenuity of experimental physicists can overcome this problem, but the second difficulty is more pernicious, and involves theory.

The Moscow group used tritium in the

form of replaced hydrogen atoms in valine, an amino acid. This was laid on a substrate forming a solid radioactive sample which could easily be inserted into a spectrometer to measure the electron energies. The drawback is that when tritium decays as part of a molecule, it excites molecular vibrations whose energy is comparable in magnitude to the putative neutrino mass. Unless these final state effects can be accurately accounted for, a reliable mass estimate cannot be retrieved from the data.

Calculations of vibrational states in complex molecules are notoriously intricate, and the uncertainties are large. This is why the Soviet claim for a neutrino mass of 30 eV or so has always been treated sceptically — the experimental technique is acknowledged to be excellent, and the results are statistically significant, but the suspicion of systematic errors has never been dispelled.

A similar experiment has been done in Zurich (M. Fritsch *et al.* *Phys. Lett.* **173B**, 485; 1986), the significant difference being that tritium was implanted in a carbon substrate. The final state effects were calculated on a basis of methane-like molecular vibrations, which ought to be simpler and more reliable than similar calculations on the complex valine molecule. The Swiss experiment yielded an upper limit of 18 eV on the neutrino mass. The contradiction between this and the equally credible Soviet results demonstrates the nature of the problem.

The ideal way to avoid all these incalculable effects is to avoid using molecules, and this is what Wilkerson *et al.* have almost done. Through ingenious design, they have been able to build an experiment to measure electrons from a sample of gaseous tritium, for which the molecular excitations are known to an accuracy of about one electron volt. The great innovation was in finding a way to transport electrons out of the chamber containing the gas so that their energy could be measured by a spectrometer. In their apparatus, tritium gas flows into the middle of a long narrow tube, and is recirculated as it passes out of the ends. The magnetic field of a superconducting solenoid forces the electrons to spiral along the tube to the ends, where they are magnetically focused and transported to a spectrometer. Use of a gaseous source also alleviates two other problems that bedevilled the earlier experiments. Where

tritium is bound in a solid sample, energy resolution suffers because electrons can be scattered within the sample itself and off the supporting substrate. In the Los Alamos experiment, internal scattering is reduced and backscattering is absent.

The results reported by Wilkerson *et al.* are derived from only twelve days of data-taking, but already they have arrived at an upper limit of 27 eV with a 95 per cent confidence level. After taking more data since their paper was submitted, they appear to be running up against some systematic errors, but with some refinement of their method believe they can get to a sensitivity of about 10 eV.

The next step will be to work with atomic rather than molecular tritium. Although the two-body final state effects are fairly well understood, they amount to a substantial correction, which it would be easier to avoid. Without molecular correction, Wilkerson says their data indicate a neutrino mass of 15 eV.

A further improvement would be to use solid instead of gaseous tritium, for an easier experimental set-up, but work in this direction at Lawrence Livermore National Laboratory has not produced results so far. The perfect tritium endpoint experiment would measure electrons from bare nuclei, so that even atomic excitations were absent. This in principle would allow errors to be reduced to fractions of an electron volt, but the necessary handling techniques do not yet exist.

This careful and time-consuming laboratory work makes an interesting contrast to the somewhat cavalier calculations of astrophysicists using the supernova data. It is undoubtedly true that the supernova neutrinos would have easily revealed a mass of 100 eV or 50 eV, but beyond that there are differences of opinion. Bahcall and Glashow (*Nature* **326**, 476; 1987) derive an upper limit of 11 eV, but Kolb, Stebbins and Turner (*Phys. Rev. D*, in the press) say that by taking into account all possible sources of error, a reliable upper limit is about 20 eV. In some sense, both results are correct, except that different levels of confidence can be put upon them. The Los Alamos experiment gives an upper limit of 27 eV with 95 per cent confidence or, if you prefer, 23 eV with 90 per cent confidence. The uncertainties and model dependence in the astrophysical arguments (as well as the paucity of data) make comparably precise statements impossible.

David Lindley

Superconducting ceramics

Oxygen defects and structure

D.O. Welch, V.J. Emery and D.E. Cox

A FULL understanding of how the properties of the new high-temperature (T_c) superconducting oxides depend on the stoichiometry and the degree of 'doping' with added impurities is probably essential to understand the mechanism of superconductivity in these materials. For example, the roles of factors such as mixed valence effects and the avoidance of instabilities associated with exactly half-filled electron bands are intimately connected with the impurity and oxygen content, and their arrangement in the crystal structure of the oxide. There have been several important recent developments along these lines, for La_2CuO_4 as well as $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$, including those of Strobel *et al.*¹, Ourmazd *et al.*² and David *et al.*³, who report their new results on pages 306, 308 and 310 of this issue, respectively.

David *et al.*³ report a determination of the structure of $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$ from high-resolution neutron powder-diffraction data. This structure has been independently determined by the same technique by at least five other groups⁴⁻⁸. The structure is orthorhombic with assigned

space group $Pmmm$ and may be viewed as a heavily-distorted, oxygen-defective, ordered derivative of a perovskite-type structure. Two-thirds of the copper atoms have four nearest-neighbour oxygens and one, more distant neighbour in a square-pyramidal arrangement, and form a two-dimensional network in the a - b planes linked through corner-shared oxygens about 0.03 nm out of the plane. The remaining one-third of the copper atoms have square-planar coordination in the b - c planes and form chains linked through corner-shared oxygens along the b -axis (see cover of this issue). Thus, low-dimensionality in the copper-oxygen structure, the role of which in high- T_c superconductivity has been debated in connection with the La_2CuO_4 -based superconductors, is also a factor here.

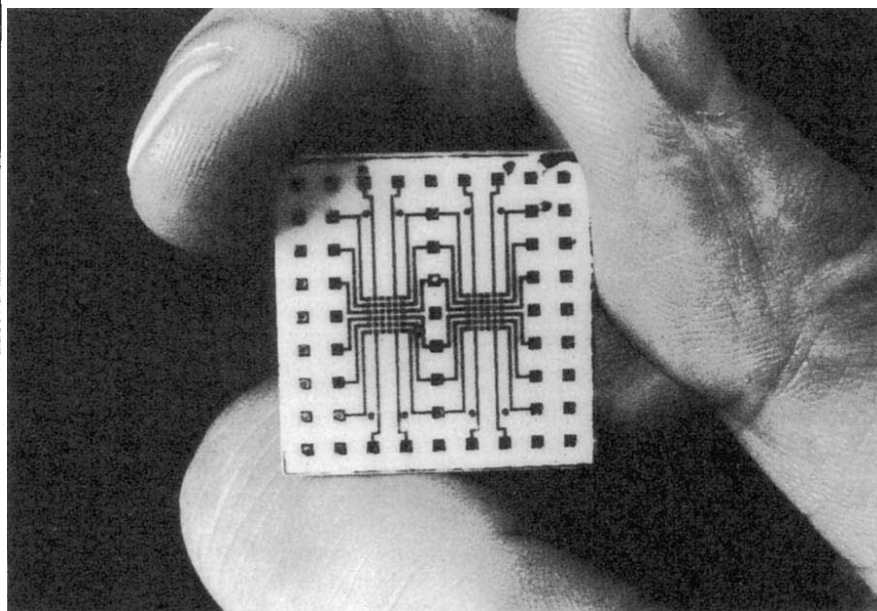
A common feature of the structures of both types of high- T_c oxide is the existence of quasi-two-dimensional CuO_2 planes. There seems to be growing evidence, however, implicating the chain-like Cu-O configuration as an important factor in the superconductivity⁹⁻¹¹. It appears that the

removal of oxygen atoms occurs most easily (and reversibly) from these chain sites, causing a substantial deterioration in T_c (even to zero as x approaches unity¹⁰ in $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$); furthermore, the tetragonal modification of the structure of this oxide may be a consequence of the destruction of the long-range chain-like structure by disordering as well as by oxygen removal⁹⁻¹¹. Strobel *et al.*¹ in their paper in this issue report a systematic study on the effect of oxygen loss, on the valence state of copper, and its relation to processing conditions. The significance of one-dimensional Cu-O structures has been pointed out by Mattheiss and Hamann (personal communication) based on their band-structure calculations for a related structure for $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$.

The one-dimensional Cu-O chain structure, inferred from neutron-diffraction data as discussed above, has also been observed directly in atomic-resolution, electron-microscope images of $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$ by Ourmazd *et al.*² in this issue, that also show numerous planar defects, which presumably result from the accommodation of the structure to deviations from ideal oxygen stoichiometry. This work clearly shows that one cannot naively assume that oxygen deficiency is simply accommodated by the creation of point defects alone. In their enthusiasm for these planar defects, Ourmazd *et al.* even make the novel suggestion that they may contribute to superconductivity by an enhancement of electron localization.

Two recent experiments on $\text{La}_2\text{CuO}_{4-y}$, related to the first high- T_c superconductors, $\text{La}_{2-x}\text{A}_x\text{CuO}_{4-y}$ ($\text{A} = \text{Ba}, \text{Sr} \dots$), are likely to have a great impact on our thinking about these materials. According to simple valence-counting arguments (and expensive electronic structure calculations) the ground state of the CuO_2 planes in La_2CuO_4 should be insulating and display some kind of long-range order: antiferromagnetism if coulombic interactions dominate; or a lattice distortion if phonon coupling is more important. Working at the High Flux Beam Reactor at Brookhaven National Laboratory, several groups found that $\text{La}_2\text{CuO}_{4-y}$ is antiferromagnetic, but only if the oxygen deficiency is greater than a few per cent¹²⁻¹⁴. At about the same time, workers from laboratories in Grenoble found that $\text{La}_2\text{CuO}_{4-y}$ is superconducting below about 38 K, provided that y is kept as small as possible (J. Beille *et al.*, personal communication).

The two results are not inconsistent as oxygen defects control the density of electrons in the CuO_2 planes. Antiferromagnetism occurs with an average of one electron per copper atom, whereas superconductivity requires a smaller number. The puzzle is that La_2CuO_4 is antiferromagnetic when non-stoichiometric in



ART meets physics: the picture shows a circuit board made by scientists at IBM's Yorktown Heights laboratory with a thin film of Y-Ba-Cu-O superconductor using a method they call spray painting. This method, in fact a well-established technique more properly called plasma spraying, involves ionizing a jet of powdered oxides in an electric arc just above the substrate surface. The plasma condenses to give a strip a few centimetres wide and a few micrometres deep which is then annealed. The method can be used to coat large and complicated surfaces with continuous films for magnetic-field shielding or to give superconducting wires as shown in the picture. The thickness of the polycrystalline film can be varied to suit the application. The device shown, made by spraying through a template, is typical of the type that would be used to give superconducting links between chips in a computer cooled in liquid nitrogen. □

oxygen, but not when stoichiometric. Simple valence arguments would suggest the opposite. It may be that there are random lanthanum defects or, alternatively, regions in which sections of the La-O planes are missing¹⁵, which alters the average carrier density in the Cu-O plane. It may also be that the electronic structure arguments are wrong.

These new observations are of great importance in that they will require a reevaluation of several models of high- T_c superconductivity. In particular, the photon-exchange mechanism¹⁶ relies on the assertion that there is a structural phase transition in La_2CuO_4 that is removed by the change in carrier density brought about by doping. This would leave 'soft' phonons which are effective in bringing about high- T_c superconductivity. Now it is unclear even that this phase transition should be sought in the undoped, stoichiometric material as this too may be superconducting. More important, antiferromagnetic ordering makes a structural

phase transition unfavourable energetically, thereby undermining the basis of high- T_c transitions by exchange of phonons. □

1. Strobel, P., Capponi, J.J., Chaillout, C., Marezio, M. & Tholence, J.L. *Nature* **327**, 306-308 (1987).
2. Ourmazd, A. *et al.* *Nature* **327**, 308-310 (1987).
3. David, W.I.F. *et al.* *Nature* **327**, 310-312 (1987).
4. Capponi, J.J. *et al.* *Europhys. Lett.* (submitted).
5. Beno, M.A. *et al.* *Appl. Phys. Lett.* (submitted).
6. Beech, F., Miraglia, S., Santoro, A. & Roth, R.S. *Phys. Rev. Lett.* (submitted).
7. Izumi, F. *et al.* *Jap. J. appl. Phys.* (submitted).
8. Cox, D.E., Moodenbaugh, A.R., Hurst, J.J. & Jones, R.H. *J. Phys. Chem. Solids* (submitted).
9. Tarascon, J.M., McKinnon, W.R., Greene, L.H., Hull, G.W. & Vogel, E.M. *Proc. 1987 Spring Meeting Mat. Res. Soc.* (in the press).
10. Nakamura, K. *et al.* *Proc. 1987 Spring Meeting of the Mat. Res. Soc.* (in the press).
11. Schuller, I.K. *et al.* *Solid St. Commun.* (submitted).
12. Vankin, D. *et al.* *Phys. Rev. Lett.* (submitted).
13. Freltoft, T. *et al.* *Phys. Rev. Lett.* (submitted).
14. Mitsuda, S. *et al.* *Phys. Rev.* (submitted).
15. Davies, A.H., & Tilley, R.J.D. *Nature* **326**, 859-861 (1987).
16. Weber, W. *Phys. Rev. Lett.* **58**, 1371-1374 (1987).

D.O. Welch, V.J. Emery and D.E. Cox are at Brookhaven National Laboratory, Upton, New York 11973, USA.

Speciation

Finding pieces of the puzzle

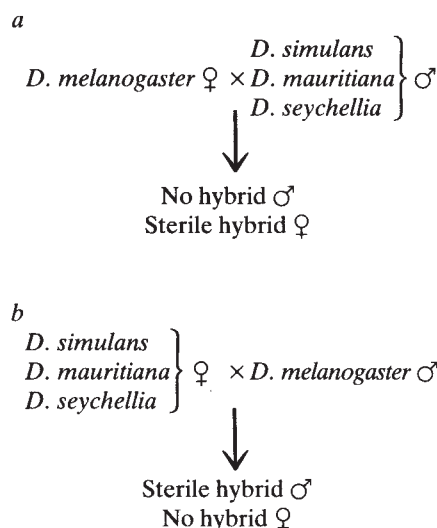
Trudy F.C. Mackay

INDIVIDUALS of a biological species are prevented from interbreeding by several pre- and post-mating reproductive isolating mechanisms. The phenomenon of speciation can be regarded as the end-product of the accumulation of genetic changes within populations caused by the microevolutionary processes of mutation, natural selection and genetic drift; and also as a basic unit for the evolution of higher taxa. What is the nature of the genetic changes that underlie species formation? Finding the answer to this question remains a challenge. The discovery by Hutter and Ashburner, reported on page 331 of this issue¹, of a gene that might be involved in the reproductive isolation of sibling species of the fruitfly *Drosophila melanogaster* species sub-group adds another tantalizing piece to the puzzle of speciation.

The difficulty that plagues experimental analysis of the genetic basis of reproductive isolation is obvious: true species do not interbreed, and genetic analysis is based on crossing. Interspecific hybridization is prevented at several stages: species differences in habitat, morphology and reproductive behaviour impede the formation of hybrids; and hybrids that do occur are often inviable, sterile or produce unfit progeny. The reduction in fitness of interspecific hybrids has been attributed to the disruption of species-specific epistatic interactions at loci controlling viability and fertility. Because it is common for related species to be separated by

more than one isolating mechanism, it has been inferred that the genetic basis of reproductive isolation must be complex. Estimates of the minimum number of loci contributing to hybrid inviability or sterility can be obtained using closely related species of *Drosophila* that are not yet fully reproductively isolated, a technique pioneered by Dobzhansky².

There are several groups of sibling species of *Drosophila* that produce viable and either partially or completely fertile hybrids. With such related species it is possible to introgress marked chromosomes or chromosome segments of one species into the genetic background of the



other, and to monitor the effect of the foreign chromosome on the expression of the trait of interest. In this way Dobzhansky² demonstrated that factors on all major chromosomes contribute to the size of the testes in backcross hybrids of *D. pseudoobscura* and *D. persimilis*, and that X-autosome interactions are important in the determination of male hybrid fertility. Subsequent analyses of the genetic basis of hybrid sterility from crosses of different *Drosophila* sibling species pairs confirm that factors on all autosomes contribute to the sterility of male-interspecific hybrids, and that species-specific epistatic interactions between factors on the X and Y chromosomes³, X and autosomes, and/or Y and autosomes⁴ are necessary for fertility.

Little is known of the genetic nature of the factors causing hybrid sterility. Only one has so far been located: a factor affecting sperm motility in *D. simulans* × *D. mauritiana* hybrids has been localized on the X chromosome to 1.1 ± 0.2 map units of the *D. simulans* forked locus⁵. Backcross hybrid males are fertile if both the Y chromosome and the segment of the X chromosome bearing the fertility factor are from *D. simulans*, but are sterile if the X-linked factor from *D. mauritiana* is paired with a *D. simulans* Y chromosome⁶. A model for the sort of interaction that might be involved is provided by the *Stellate* locus, a tandemly repeated sequence of 1,250 base pairs present in variable copy number on the X chromosome of *D. melanogaster* (the *Ste*⁺ allele has a low copy number of the sequence, whereas the *Ste* allele has about 200 copies), and which is also present in about 80 copies on the *D. melanogaster* Y chromosome⁷. Males with deleted copies on the Y chromosome are sterile, and it has been proposed that the gene product of the *Stellate* copies on the Y chromosome normally represses transcription of *Stellate* genes on both the X and Y chromosomes (ref. 7). Derepression would then result in overproduction of the *Stellate* protein by the X-linked copies, which might accumulate to form the characteristic crystals in the primary spermatocytes of males lacking the *Stellate* control region. The *Stellate* sequence is present in reduced copy number and only on the Y chromosome of *D. simulans* and *D. mauritiana*; it is completely absent from *D. erecta*, *D. tiansieri* and *D. yakuba*. If other species-specific fertility factors interact in a manner analogous to *Stellate*, hybrid sterility could well result from disruption in regulation caused by removing one or other of the interacting components.

If little is known of the genetic basis of hybrid sterility, even less is known about hybrid viability. One group of sibling species for which the reproductive relationships have been well-characterized are four species of the *melanogaster*

group: *D. melanogaster*, *D. simulans*, *D. mauritiana* and *D. seychellia*. Crosses of *D. melanogaster* females to males of any of the other three species give sterile female hybrids and males that die as third-instar larvae, (*a* in the figure) whereas crosses of females of the sibling species to *D. melanogaster* give sterile male hybrids and females that do not survive embryogenesis⁸ (*b* in the figure). Hence it seems that incompatibilities between a factor or factors on the *melanogaster* X and *simulans* Y chromosomes, *melanogaster* X and autosomes, and/or *simulans* Y and autosomes are implicated in male hybrid survival. Female survival seems to be determined by an interaction between genotype and a component of maternal cytoplasm, as genotypically identical female hybrids from the reciprocal crosses are viable only if the cytoplasm is from *D. melanogaster*.

Some progress towards understanding the genetic basis of interspecific hybrid inviability in crosses of *D. melanogaster* to its sibling species may come from analysis of mutant alleles at loci with the peculiar property that they 'rescue' the inviable sex. The first of these, *Lhr* (Lethal hybrid rescue) was discovered⁹ in *D. simulans* and mapped to chromosome 2R; when *D. melanogaster* females are crossed to *Lhr* *D. simulans* males, male progeny survive. Hutter and Ashburner in this issue¹ report the screening of 63 *D. melanogaster* populations for an analogous mutation. They find one population that is polymorphic for a mutant allele which they call *Hmr* (Hybrid male rescue), and which they map to position 1–31.84 of the X chromosome. When *Hmr* *D. melanogaster*

females are crossed to any strain of *D. simulans*, survival of hybrid male progeny nearly equals female survival at 18 °C. *Hmr* rescue is temperature sensitive, with the critical period for rescue at mid-to-late second-instar larva. *Hmr* also rescues male hybrids from interspecific crosses of *D. mauritiana* and *D. seychellia* males to *Hmr* *D. melanogaster* females, but has no effect on female survival in the reciprocal cross. There is, however, a strain of *D. melanogaster* whose males rescue hybrid females from the reciprocal cross to *D. simulans* or *D. mauritiana* females⁸. The factor (or factors) responsible for hybrid female rescue have not yet been analysed.

What is the function of the *Hmr* locus in *D. melanogaster*? Is it related to the *Lhr* locus in *D. simulans* or to the unidentified factor that rescues hybrid females? How does *Hmr* interact with other loci? *Drosophila* genetic technology is now sufficiently sophisticated that cloning *Hmr* is a feasible goal, allowing experiments designed to answer these and related questions. The opportunity to study one of the genes involved in species formation is an exciting prospect. □

1. Hutter, P. & Ashburner, M. *Nature* **327**, 331–333 (1987).
2. Dobzhansky, T. *Genetics* **21**, 113–135 (1936).
3. Coyne, J.A. *Nature* **314**, 736–738 (1985).
4. Vigneault, G. & Zouros, E. *Evolution* **40**, 1160–1170 (1986).
5. Coyne, J.A. & Charlesworth, B. *Heredity* **57**, 243–246 (1986).
6. Coyne, J.A. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4444–4447 (1984).
7. Livak, K.J. *Genetics* **107**, 611–634 (1984).
8. Lachaise, D., David, J.R., Lemeunier, F., Tsacas, L. & Ashburner, M. *Evolution* **40**, 262–271 (1986).
9. Watanabe, T.K. *Jap. J. Genet.* **54**, 325–331 (1979).

Trudy F.C. Mackay is in the Department of Genetics, University of Edinburgh, West Mains Road, Edinburgh EH9 3JN, UK.

preferential occurrence of these sequences in nucleosomal DNA is also periodically modulated. For uniform bending, in which the direction of curvature is constant, the sequence periodicity and the double-helical periodicity should be the same^{4,5}. This means that the particular sequences correlated with either an inward- or an outward-facing minor groove tend to occur in the nucleosome at intervals of approximately 10 base pairs. As we discuss below, this 'rule of 10' is also apparent in DNA sequences involved in other protein–DNA complexes, implying that here too the DNA is bent into a nucleosome-like configuration.

The nucleosome is an example of a nucleoprotein complex in which DNA bendability is an important determinant of

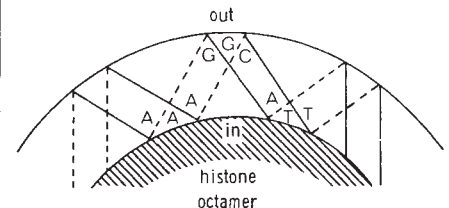


Fig. 1 Preferred orientation of three trinucleotide sequences as deduced for the nucleosome core particle¹³. Parallel lines, DNA minor groove; lettering, inferred orientation of the minor grooves of the three trinucleotides relative to the histone octamer. Note that the inferred positions are an average of the sequence shown and of its complement. Thus, for example, AAA represents the average position of itself and of its complement, TTT.

specific positioning. In many other such complexes, however, DNA bending seems to complement sequence-specific recognition. The relative contribution of bending to complex formation must depend both on the magnitude of curvature and on the length of the curved DNA segment. Because it seems probable that the energy required to deform DNA mildly in a non-preferred direction is a good deal less than that which might be available for direct sequence recognition (by, for example, hydrogen bonding) the influence of bendability must, in general, be a second-order effect.

One well-studied case of complex formation involving both DNA bending and sequence-specific recognition is the interaction of the *Escherichia coli* catabolite activator protein (CAP) with the *lac* control region. The existence of significant DNA bending has been inferred from the anomalous mobility of the complex in polyacrylamide gels⁶; from the effect of CAP on the rate of ligation of small circular DNA molecules containing the *lac* control region⁷; and also from electron micrographs of the complex⁸. These studies provide little direct indication of the magnitude of curvature or of the orientation of the bent DNA, that is, whether it is bent towards or away from the CAP dimer. Recent results from

Nucleoprotein complexes

DNA wrapping and writhing

Andrew Travers and Aaron Klug

In many DNA–protein complexes the DNA is wrapped tightly around a central core of protein. The archetypal example of such a complex is the nucleosome core particle, in which about 145 base pairs (bp) of DNA are wound more or less uniformly in approximately 1.8 superhelical turns around the histone octamer¹. The diameter of this superhelix is 86 Å, only four times that of the DNA double helix itself. To accommodate curvature of this magnitude all the grooves (both major and minor) on the inside of the curve must narrow considerably, owing to the compression associated with bending, whereas those on the outside of the curve become correspondingly wider. The changes in the base-pair stacking necessary to achieve this will depend on the rotational setting of the DNA relative to the direction of bending. In a recent News and Views

article², T.J. Richmond discussed some aspects of kinking in DNA–protein interactions, and here we will discuss the large-scale changes in DNA configuration (writhing) found in other nucleoprotein complexes.

The tight bending of DNA by a protein requires that particular stacking interactions between adjacent base pairs can accommodate the deformations of groove width. In nucleosomal DNA the variation in groove width is correlated with the preferred occurrence of particular short sequences such that, for example, the trinucleotide AAA or TTT is found more often where the minor groove points in towards the histone octamer, whereas another trinucleotide GGC or GCC prefers to be sited in an outward-facing minor groove³ (Fig. 1). Because the variation of minor-groove width is periodic, the

Crothers' laboratory⁹, however, show that in this complex strong binding interactions span 28–30 bp of DNA and that the magnitude of bending is affected by mutations just outside this region. From the dimensions of the CAP dimers and the assumed energetics of complex formation¹⁰, Liu-Johnson *et al.*⁹ propose that the magnitude of curvature in the CAP–DNA complex is similar to that in the nucleosome core particle and that this curvature is directed towards the protein.

As the CAP binding site contains an interrupted inverted repeat with major-groove contacts half a double-helical turn on each side of the centre of symmetry, the dyad at the centre must have its minor groove facing inward towards the protein (Fig. 2). This assignment of the direction of curvature can be brought⁵ into striking agreement with the sequence of the *lac* CAP binding site using the rule of 10 deduced from DNA bending on the nucleosome. At the centre of dyad symmetry and at 10 bp on either side are AT-rich sequences which on the nucleosome occur preferentially where the minor groove is narrowed. Similarly exactly out of phase are sequences that preferentially occupy an outward-facing minor groove. Two conclusions can be drawn from this precise correspondence between the CAP binding site and nucleosomal DNA. First, it seems probable that DNA bendability makes a significant contribution to the stability of the CAP–DNA complex. Indeed, separate alterations to two of the short sequences in the CAP binding site correlated with bendability both result in a weaker complex in which the bending is significantly less, although it is not possible from the available evidence to exclude alteration of specific contacts. The second conclusion is that the sequences involved in specific recognition at the major-groove contacts, that is, GTG and CAC, are also compatible with the direction of bending in the complex.

A second example of a highly conserved sequence compatible with the curvature of DNA on the protein surface is found in the binding site of *E. coli* RNA polymerase. As with CAP, the DNA is bent around the protein¹¹ and the sequences in the promoter site exhibit similar alternations of

preferred 'bending' sequences to those found in the nucleosome. Within the very highly conserved –10 region the sequence AAT is located at an inward-facing minor groove as on the nucleosome and is separated by half a double-helical turn from preferred outward-facing sequences. Again, this result is consistent with dual roles for short sequences in nucleoprotein complexes.

The formation of nucleosome-like structures is a common feature of processes such as site-specific recombination and the initiation of DNA replication that involve the enzymatic manipulation of DNA. These structures can, as in the cases of DNA gyrase or Tn3 resolvase, bring into close proximity DNA sequences that would otherwise be spatially distant from each other¹². A simple example of such a complex is provided by the resolvase encoded by the Tn3, or closely related $\gamma\delta$, transposon, a dimeric protein which binds to three neighbouring sites on each of the recombining DNA molecules^{13–15}. Two of these sites, II and III, follow the

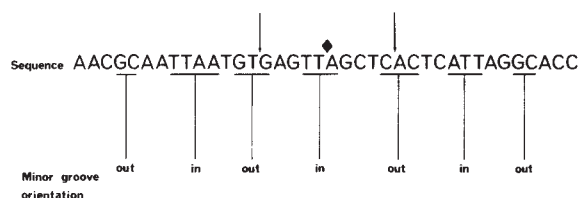


Fig. 2 Sequence of the CAP binding site in the *E. coli lac* promoter region showing the relationship of the DNA sequence to the orientation of the minor groove. Arrows, position of the DNase I-accessible internucleotide bonds, which must therefore be located where the minor groove faces away from the protein²⁴. Underlined sequences, tri- or tetranucleotide sequences which, on the nucleosome, would preferentially occupy inward- or outward-facing minor grooves.

The diamond represents the binding-site dyad.

rule of 10 over approximately three and two double-helical turns, respectively, between the close contact points (M. Boocock, J.L. Brown and D. J. Sherratt, personal communication). Site I, which contains the actual site of recombination, is however anomalous and does not observe this rule, having a separation of about 2.3 turns between these same close contact points. If the dimeric resolvase protein were to make the same homologous contacts with its recognition sequence

at all sites then either the rotation of the individual subunits in the dimer relative to each other, or the average twist of the DNA between the contact points in site I, must change. This latter possibility seems more probable, as the relatively unstable TATA sequence, which could readily locally unwind^{16,17} (or possibly kink¹⁸), is at the crossover point. This would alter the average twist at site I and would at the same time maintain the relationship between the number of double-helical turns and the direction of bending at all three binding sites. Such an unwinding would also allow site I to be functionally distinguished from the other two sites.

The mechanics of recombination require that the two recombining DNA helices be closely juxtaposed in a precise spatial arrangement. This arrangement can determine the topological and directional outcome of the recombination event^{19,20}. One example of directional site-specific recombination is the integration of phage λ DNA into the bacterial

chromosome. This recombination takes place between two attachment sites on the bacterial and phage chromosomes, *attB* and *attP*, respectively. These sites differ substantially in complexity, *attB* being 25 bp long and *attP*

240 bp long and wrapped in a nucleosome-like structure²¹, the intasome, around a protein complex containing the λ *int* protein and the integration host factor. Recombination requires both an intact *attP* site and negative supercoiling of *attP* but not

attB DNA. Using an indirect end-labelling footprinting technique, Richet *et al.*²² have recently demonstrated that negative supercoiling enhances the affinity of *int* protein for intact but not truncated *attP*. The extent of the protected region, and therefore of the DNA covered by protein, however, is the same in both cases. Richet *et al.*²² therefore suggest that the primary role of supercoiling is to facilitate the formation of a wrapped, functionally active, higher-order structure rather like the nucleosome. A similar situation has been observed in the assembly of nucleosomes²³, implying that the DNA in the intasome is wrapped in a left-handed superhelix. In these assembled structures there is no evidence for structural distortion of the DNA at the crossover point in *attP*, and consequently Richet *et al.* suggest that the role of the intasome is to facilitate precise pairing interactions of *attP* with *attB* and thus drive the recombination event in a unique direction. □

Andrew Travers and Aaron Klug are at the MRC Laboratory of Molecular Biology, Cambridge CB2 2QH, UK.

- Richmond, T.J. *et al.* *Nature* **311**, 532–537 (1984).
- Richmond, T.J. *Nature* **326**, 18–19 (1987).
- Satchwell, S.C., Drew, H.R. & Travers, A.A. *J. molec. Biol.* **191**, 659–675 (1986).
- Drew, H.R. & Travers, A.A. *J. molec. Biol.* **186**, 773–790 (1985).
- Travers, A.A. & Klug, A. *Phil. Trans. R. Soc.* (in the press).
- Wu, H.-M. & Crothers, D.M. *Nature* **308**, 509–513 (1984).
- Kotlarz, D., Fritsch, A. & Buc, H. *EMBO J.* **5**, 799–803 (1986).
- Gronenborn, A.M. *et al.* *J. molec. Biol.* **179**, 751–757 (1984).
- Liu-Johnson, H.N., Gartenberg, M.R. & Crothers, D.M. *Cell* **47**, 995–1005 (1986).
- Weber, I.T. & Steitz, T.A. *Proc. natn. Acad. Sci. U.S.A.* **81**, 3973–3977 (1984).
- Kunke, G., Fritz, H.J. & Ehring, R. *EMBO J.* **6**, 507–513 (1987).
- Echols, H. *Science* **233**, 1050–1056 (1986).
- Grindley, N.D.F. *et al.* *Cell* **32**, 19–27 (1982).

- Sherratt, D.J. *et al.* *Cold Spring Harb. Symp. quant. Biol.* **59**, 227–233 (1984).
- Falvey, E. & Grindley, N.D.F. *EMBO J.* **6**, 815–821 (1987).
- Drew, H.R., Weeks, J.R. & Travers, A.A. *EMBO J.* **4**, 1025–1032 (1985).
- McClellan, J.A., Palacek, E. & Lilley, D.M.J. *Nucleic Acids Res.* **14**, 9291–9309 (1986).
- Hatfull, G.F., Noble, S.Z. & Grindley, N.D.F. *Cell* **49**, 103–110 (1987).
- Wasserman, S.A. & Cozzarelli, N.R. *Science* **232**, 951–960 (1986).
- Sadowski, P. *J. Bact.* **165**, 341–347 (1986).
- Better, M. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **79**, 5837–5840 (1982).
- Richet, E., Abcarian, P. & Nash, H.A. *Cell* **46**, 1011–1021 (1986).
- Germond, J.E. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **72**, 1843–1847 (1975).
- Spassky, A., Busby, S. & Buc, H. *EMBO J.* **3**, 43–50 (1984).

Phenylketonuria

Population genetics of a disease

Kenneth K. Kidd

PHENYLKETONURIA (PKU), one of the commonest recessively inherited genetic diseases, occurs in northern European populations in roughly 1 per 10,000 live births. It is caused by a deficiency of the enzyme phenylalanine hydroxylase (PAH) and results in mental deficiency unless the affected child is kept on a strict, expensive diet for at least the first decade of life, and perhaps for life. DiLella and colleagues last year performed¹ a molecular-genetic analysis of a large sample of families and found that for a surprisingly high proportion of cases the abnormal PKU alleles had a common origin derived from a single mutational event. They have now extended this analysis and on page 333 of this issue² report that just two mutations account for as many as 50 per cent of cases in a sample of northern European origin. Similar analyses on Yemenite Jewish populations have begun to reveal a similarly limited number of mutations. This is surprising in a genetic disease as common as phenylketonuria and raises the question of whether the PKU allele has a selective advantage in heterozygotes.

Insight

These insights into the population-genetic and evolutionary history of PKU have been made possible by the use of restriction-fragment length polymorphisms (RFLPs) to identify the origins of the mutation in the PAH gene that is the cause of the disease. RFLPs reflect the normal variation in the DNA around the PAH locus³; eight different RFLPs have been defined using the PAH complementary DNA as the probe⁴. Because these eight RFLPs can occur in different combinations on different chromosomes in the population, there is a total of 256 possible patterns of identifiable variation (haplotypes) in this region of the chromosome. Only a small subset of these haplotypes is actually seen, but in most people the haplotypes at the PAH locus are distinguishable and, therefore, it is easy to follow the transmission of specific haplotypes through families. The population frequencies of the haplotypes are the results of the evolutionary history of the PAH locus — the specific haplotype on which a new RFLP mutant occurred; random factors affecting haplotype frequencies in the population; rare recombination events within the locus creating new haplotypes; and whatever selective forces might be operating on the PAH locus or on nearby loci. Whatever that history, the result is very strong linkage

disequilibrium among the eight RFLPs, that is, several haplotypes occur in frequencies far greater than expected from a random combination of the specific DNA sequences at the RFLP sites. Several distinguishable haplotypes carry a PKU mutation, but these PKU haplotypes occur in frequencies very different than for normal PAH haplotypes⁴.

Because a mutational event creating a PKU allele (a PAH allele producing little or no functional enzyme) must occur in a chromosome already characterized by some specific PAH haplotype, the PKU chromosomes that are direct descendants of a single mutation will all tend to have the same RFLP haplotype. In general, only two factors will cause mutants to be distributed at random among the haplotypes seen on normal chromosomes: frequent re-occurrence of mutations or sufficient evolutionary time (with respect to the recombination rate between the mutant and the RFLPs characterizing the haplotypes) to shuffle the mutation among the haplotypes. Thus, distributions of PAH haplotypes might be expected to be different in normal and PKU chromosomes, that is, to show linkage disequilibrium. Similar findings occur at other loci. For example, only three of more than a dozen different RFLP haplotypes at the β -haemoglobin cluster contain the haemoglobin-S mutation^{5,6}.

This strong linkage disequilibrium in PKU alleles and PAH RFLP haplotypes has been used by DiLella and associates to identify and characterize specific mutations causing PKU. They reasoned that PKU patients homozygous for a PAH haplotype are likely to carry the same mutation in the PAH gene; this was exactly the logic used to clone and characterize β -thalassaemia mutants⁷. In their first paper¹, the authors demonstrated very strong linkage disequilibrium between a specific splice-junction mutation and a specific RFLP haplotype. This haplotype (No. 3) occurs with a frequency of only 3 per cent among normal (non-PKU) chromosomes but accounts for 38 per cent of the PKU chromosomes. Using a pair of synthetic oligonucleotides (one for the normal and one for the mutant sequence) they showed that this splice-junction mutation occurs on all haplotype-3 PKU chromosomes and never (in their sample) on other PKU chromosomes. The obvious conclusion is that a single mutational event altered a splice-junction sequence of PAH in a haplotype-3 chromosome. All current haplotype-3 PKU alleles are evidently descendants of this

single mutated chromosome and now account for more than one-third of PKU alleles in the Danish population.

Having proved their logic correct, DiLella *et al.* report in this issue² the cloning, characterization and population genetics for a second PKU mutation. They cloned the PKU allele from a patient homozygous for haplotype-2, the second most common PKU allele, and show the mutation to be an amino-acid substitution. They show, with a pair of synthetic oligonucleotides to the normal and mutant sequences, that this particular mutation is in linkage disequilibrium with haplotype-2. Again, the obvious conclusion is that a specific mutational event occurred in a normal haplotype-2 PAH gene and that all current haplotype-2 PKU alleles are direct descendants of that single mutant. These descendant PKU alleles now account for 20 per cent of PKU alleles.

Screening

The two mutants now characterized account for more than 50 per cent of all PKU alleles in the northern European sample. There is every reason to expect that the next two PKU haplotypes to be characterized will each be characteristic of a different PKU mutant and that these will also be converted by synthetic oligonucleotides into codominant genetic systems. If this expectation materializes, it has important implications both clinically and in population genetics. Population screening is already a real possibility, as more than half of all PKU carriers can be recognized using relatively simple tests with the two pairs of synthetic oligonucleotides; that proportion may be increased to 90 per cent when the associated mutations of the other two haplotypes are characterized. The frequency of unidentified carriers would then be only one-tenth the current frequency of carriers and the frequency of PKU children born to two parents who are not detectably carriers would be 1 per cent of the present incidence of PKU. The evolutionary and population-genetic consequences of these findings over the next few generations will be determined by how society and individual couples choose to deal with the new information potentially available to them. For example, the frequency of the disease would be dramatically reduced were known carriers to avoid having children with each other.

Whatever the evolutionary future, the new findings help to elucidate the evolutionary past of the PAH locus. First, the hypothesis that a high mutation rate accounts for the high frequency of PKU is clearly excluded, leaving random genetic drift or natural selection as the only possible explanations for the high frequency of PKU in northern European populations. The data seem to argue against random

genetic drift. Although drift could cause one mutant allele to increase to a high frequency, it seems improbable that drift would cause at least two, and probably four, distinct mutant chromosomes all to increase in frequency. Moreover, the multiple distinct mutants at high frequency argue against a simple hitch-hiking hypothesis in which selection at a nearby locus elevates the frequency of a new mutant. Might there be or have been some selective advantage to individuals heterozygous for a PKU allele? It is already well known that heterozygotes for the hereditary anaemias are resistant to malaria and this accounts for the high incidence of these mutations in areas where malaria is endemic. What might such an advantage be for heterozygotes for PKU? There are no clear answers; but some type of selection for PKU alleles seems possible. The possibility that selection operates at this locus, even among normal PAH alleles, is suggested by the very nature of the two different haplotypes (1 and 4). These two haplotypes occur at almost exactly equal frequencies. For the five RFLPs in the 30 kilobases at the 3' end of the PAH locus, these haplotypes are complementary — where one has a restriction site present the other does not⁴. This observation suggests that, at least for this part of the gene, these two haplotypes have been evolving separately for a long time. Although random genetic drift could cause almost any distribution, the equal frequencies of these ancient haplotypes is very suggestive of balancing selection.

Information on the frequencies, in other populations, of RFLP haplotypes for both normal and PKU chromosomes may help resolve these and other questions, such as the origin of PKU in the Yemenite Jews⁵: was the PKU allele brought in from Europe or did it arise independently? In the past few weeks S. Avigad, B.E. Cohen and Y. Shiloh at the Sackler School of Medicine, Tel Aviv University, have characterized the PKU mutant in the Yemenite Jews (personal communication). These unpublished data indicate that in five different families all the PKU alleles have a deletion of an entire exon. Shiloh and colleagues have fully characterized the two PKU chromosomes in one patient and find that both are deletion alleles and have the same RFLP haplotype, one that has not been observed in European caucasians. Based on their preliminary and still partial characterizations, the haplotypes in the remaining families seem to be the same. Again, a single mutation (deletion) seems to have occurred and to have increased in frequency along with the haplotype in which it occurred. This PKU allele is not of European origin, but either drift or selection could explain the high frequency of PKU alleles in Yemenite Jews.

Louis de Broglie (1892–1987)

LOUIS de Broglie, who died on 19 March, was the last surviving, great founder of quantum physics. His theoretical discovery, in 1923, of the existence of matter-waves was the starting point of all the developments based on the notion of the wavefunction, from which Schrödinger derived the equation bearing his name. Louis de Broglie's ideas were thus the direct source of the new mechanics that now pervades the whole of physics and that came to be known as quantum mechanics after it was formally identified with the mechanics independently developed in 1925–26 in Copenhagen and Göttingen. In 1929, de Broglie was awarded the Nobel prize for this work.

The circumstances that led Louis de Broglie to his basic idea are interesting. Maurice de Broglie, his elder brother, was a physicist himself, and had a private laboratory in which he performed significant experiments on X-rays. Louis first studied history for a short time, but when Maurice showed him the texts of the 1911 1st Solvay Conference (of which he had been a scientific secretary) Louis, then aged 19, was so interested by the new problems discussed there that he turned to theoretical physics and particularly to analytical mechanics. After World War I he resumed his studies in the field and in 1923 he published three notes in the *Comptes Rendus de l'Académie des Sciences*.

These three notes are basic. In the first, Louis de Broglie proposed to associate a wave with any free particle and, making use of relativistic arguments, he derived the value of the wavelength (this is the famous Louis de Broglie formula). In the third, he demonstrated the equivalence of the principle of least action applied to a particle and Fermat's principle applied to the associated matter-wave.

By the end of 1923 Louis de Broglie had already mastered some of the main concepts of what was to become quantum physics and during 1924 he developed his ideas in his PhD thesis. This was sent by Paul Langevin to Einstein, who immediately understood its fundamental importance, drew the attention of the scientific community to it and used the

ideas in it in his development of Bose–Einstein statistics.

In the course of his life, Louis de Broglie made a number of other successful investigations. He was, for example, the first to write down the relativistic equations for a massive, spin-1, free particle in 1934, and during the Second World War he constructed the equations describing the motion of spin- $n/2$ particles, n being any positive integer. Throughout his life, however, Louis de Broglie remained puzzled, at a deep level, by the discovery that he himself had started in 1923.

Although quantum mechanics worked so beautifully, he felt, as did Einstein, that neither he nor anyone else understood what it meant as a description of nature. This belief soon prompted him to try to interpret the quantum-mechanical formalism by considering a particle as a singularity in a wave (the 'double solution' theory). He presented the simplified 'pilot-wave' variant of this theory at the 5th Solvay Conference in 1927. It is the ancestor of all the so-called 'hidden-variables' interpretations of quantum physics. In fact the main ideas and formulas of David Bohm's hidden-variable theory, published in 1952, were essentially rediscoveries of this 25-year-old pilot-wave theory (although Bohm's work did go much further). During his last years of study Louis de Broglie returned again to the problem of a proper understanding of quantum physics.

Louis de Broglie was a polite and unaffected man. He belonged to a very aristocratic family (he inherited the title of Duke from his brother) but his interests were almost exclusively directed towards scientific knowledge. Despite all his illustrious charges (Académie des Sciences, Académie Française) he led a very simple, almost ascetic life. He was a professor at the Sorbonne and he also wrote excellent, popular works describing to a large audience the conceptual problems raised by twentieth century physics. Bernard d'Espagnat

Bernard d'Espagnat is at the Laboratoire de Physique Théorique et Particules Elementaire, University of Paris, 91405 Orsay Cedex, France.

The studies I have discussed here are changing the way medical geneticists must deal with counselling for PKU; they offer all interested individuals the ability to have much more certain information on their carrier status. The studies also illustrate both the complexity of the evolutionary history of a major genetic disease and the great power of molecular-genetic techniques to illuminate the evolutionary facts, if not yet the evolutionary forces. □

I. DiLella, A. G., Marvit, J., Lidsky, A. S., Guttler, F. & Woo, S. L. C. *Nature* 322, 799–803 (1986).

- DiLella, A. G., Marvit, J., Brayton, K. & Woo, S. L. C. *Nature* 327, 333–336 (1987).
- Woo, S. L. C., Lidsky, A. S., Guttler, F., Chandra, T. & Robson, K. J. H. *Nature* 306, 151–155 (1983).
- Chakraborty, R. *et al. Hum. Genet.* (in the press).
- Antonarakis, S. E. *et al. Proc. natn. Acad. Sci. U.S.A.* 81, 853–856 (1984).
- Pagnier, J. *et al. Proc. natn. Acad. Sci. U.S.A.* 81, 1771–1773 (1984).
- Orkin, S. H. & Kazazian, H. H. Jr *A. Rev. Genet.* 18, 131–171 (1984).
- Adam, A. *Israel J. Med. Sci.* 9, 1383–1392 (1973).

Kenneth K. Kidd is in the Department of Human Genetics, Yale University School of Medicine, PO Box 3333, New Haven, Connecticut 06510-8005, USA.

Photosynthesis

Plant reaction centre defined

M. C. W. Evans

VIRTUALLY all the energy used by living organisms comes from the Sun. Light energy is converted to chemical energy by photosynthesis, a process occurring in the photochemical reaction centre of plants, algae and photosynthetic bacteria. There are two photosystems in plants, and Namba and Satoh (ref. 1 and data presented at a Japan-US collaboration seminar, Okazaki, 17-22 March 1987) now seem to have elucidated the structure of the complex forming the basis of the photosystem 2 reaction centre.

The initial event in photosynthesis is the photochemical oxidation of chlorophyll and coupled donation of an electron to a low-potential electron acceptor. Subsequent transfer of the electron through the electron-transport chain leads to ATP synthesis and reduction of NADP. These are then used for CO₂ fixation and other biosynthetic processes.

The photochemical oxidation of chlorophyll and reduction of an electron acceptor can be observed in solution, but this is an extremely inefficient process that cannot trap useful amounts of energy. The inefficiency is partly caused by the very short lifetime of the excited state of chlorophyll of about 20 ns and partly by back-reactions between oxidized and reduced products. Photosynthetic organisms overcome this inefficiency by incorporating the chlorophyll and acceptor into a lipoprotein complex, a reaction centre, and incorporating this complex into a membrane with specific orientation to separate the oxidized and reduced products of the photochemical reaction.

The concept of a specific photochemical reaction centre developed from the work of Duysens², who originally showed that only a small fraction of the chlorophyll in purple photosynthetic bacteria is photo-oxidized and that the active chlorophyll (P890) has a longer wavelength-absorption peak than the bulk chlorophyll. This indicated a slightly lower excitation energy for the reaction centre, allowing it to act as an energy sink, collecting energy initially absorbed by the bulk of the chlorophyll. This bulk chlorophyll is associated with specific proteins forming a light-collecting complex to increase the absorption cross-section of the reaction centre. Reaction-centre chlorophylls were subsequently identified in the two photosystems of plants, P700 in photosystem 1 and P680 in photosystem 2.

Clayton and co-workers then developed procedures for the isolation of reaction centres from the purple photosynthetic bacteria³, which contain three poly-

peptides, L, M and H, four bacteriochlorophylls, two bacteriopheophytins, one or two quinones and an iron atom. (Reaction centres from some species also contain bound cytochromes.) Spectroscopic studies of reaction centres show that the initial photochemical event involves the transfer of an electron from a pair of bacteriochlorophyll molecules to one of the bacteriopheophytin molecules in a few picoseconds. The subsequent transfer of the electron to a quinone occurs in about 150 ps, stabilizing the charge separation into the time domain of normal chemical reactions.

In one of the most exciting developments in photosynthesis and in membrane biology, Michel, Dessenhofer and colleagues crystallized the reaction centre from *Rhodospseudomonas viridis*⁴ and



Stereo view of the cytochrome subunit. The protein chain is a ribbon; haem groups and cysteine side chains are line drawings. (From ref. 5.)

elucidated the structure by X-ray crystallography⁵, at the same time determining the amino-acid sequences of the polypeptides, providing detailed information about the structural relationships of the prosthetic groups (see figure).

The isolation of reaction centres from plants and cyanobacteria has been less successful until very recently. Photosystem 2 is particularly difficult, perhaps because of the problem of assaying P680 or any other reliable indicator of activity. Despite this problem, photosystem 2 has attracted much interest because of its role as the source of the oxidant for water and because of the growing analogy to the well-understood purple bacterial reaction centre. The stable acceptor of photosystem 2 is known to be a quinone and, following the characterization of the bacterial system, it was shown that a pheophytin is the intermediary electron carrier and that magnetic interactions between the quinone pheophytin and an iron atom have a very close analogy to the purple

bacterial electron-acceptor complex.

Identification of the polypeptides forming the physical basis of the reaction centre has been more difficult. Isolation of oxygen-evolving preparations improved to the point where all the activity of the starting material can be retained with complete separation of photosystems 1 and 2. But the photosystem 2 preparations are very complex, containing all the peptides associated with oxygen evolution and all the light-collecting chlorophyll proteins. 'Core' complexes isolated from these preparations contain five polypeptides, two of which, CP43 and CP47, are chlorophyll proteins. Another, D1, of relative molecular mass 34,000 (34K) was identified as a quinone-binding protein on the basis of its ability to bind herbicides such as atrazine and its involvement in herbicide resistance. A fourth polypeptide, of mass about 31K and called D2, has no known function, although mutants of *Scenedesmus* lacking this protein have no photosystem 2 (ref. 6). The last polypeptide, of 8K, is the apoprotein of cytochrome *b*₅₅₉. Attempts to identify the reaction centre concentrated on the two chlorophyll proteins. The difficulty of measuring photosystem 2 activity led to the use of the fluorescence characteristics, particularly the emission of fluorescence at 695 nm which was thought to arise from emission following the back-reaction between P680⁺ and reduced pheophytin, and was shown⁷ to be associated with CP47. Removal of CP43 by gene deletion⁸ resulted in loss of photosystem 2 but no activity could be associated with the isolated protein, so the case for the CP47 seemed very strong.

As this work progressed, an alternative approach developed as the amino-acid sequences of the L and M subunits of the bacterial reaction centre and then of D1 and D2 became available. The sequences of all four polypeptides are very similar, which led Trebst⁹ to propose that the photosystem 2 reaction centre would be a D1/D2 complex analogous to the L/M complex. This proposal, strongly supported by Michel, offers a novel approach to the isolation of the reaction centre. Instead of looking for activity, or even chlorophyll, purification of D1/D2 can be followed, and this is the approach that Osamu Namba and Kimiyuki Satoh (Okayama University) have used successfully.

At the International Congress of Photosynthesis Research in Providence, Rhode Island last summer these authors presented preliminary evidence that they had obtained a D1/D2 cytochrome *b*₅₅₉ preparation with 4-6 chlorophyll and 2 pheophytin molecules¹. This complex had one photochemical activity, the formation of a spin-polarized triplet at low temperatures, characteristic of all reaction centres. This triplet arises from the reaction-centre chlorophyll but forms only as a result of

electron transfer between the reduced pheophytin and the oxidized reaction centre, and is therefore diagnostic for reaction-centre function. The properties of the triplet vary slightly for different reaction centres, allowing each one to be identified. The triplet observed in the preparation had the properties of the photosystem 2 triplet. The preparation would accumulate reduced pheophytin if illuminated under reducing conditions, and this is also seen in less-purified preparations. At that time it was clear that extensive characterization of the preparation would be required before it could be accepted as the definitive photosystem 2 reaction centre. D1 and D2 were identified only as stained bands on SDS gels; detection of P680, kinetic analysis of the photochemistry and an explanation of the CP47 results were all required. The last explanation was particularly important as at the Rhode Island meeting, Kazuhiko Satoh (Tokyo University) described a preparation from a cyanobacterium that has photosystem 2 activity and apparently lacks D1 and D2 (ref. 10).

Since then, intensive work has provided much of the required information, which was presented at the Okazaki meeting this year. The identification of D1 and D2 has been confirmed by the use of antibodies. Kimiyuki Satoh used antibodies against the isolated peptides and against synthetic peptide fragments made by Trebst's group. J. Barber and co-workers (Imperial College, London) isolated the preparation independently and used antibodies against the D1 and D2 gene products cloned in *Escherichia coli* to identify the polypeptides. They showed that a 60K band on gels is an unresolved D1/D2 complex reacting with both antibodies, which turns out to be the explanation of Kazuhiko Satoh's apparently D1/D2-free preparation. He now finds his preparation contains a complex reacting with both antibodies that is not resolved on gels.

Absorption spectra of the preparation show the chlorophyll and pheophytin. The pheophytin peak is split, suggesting slightly different environments for the two molecules. Kinetic optical measurements of flash-induced absorption changes in H. van Gorkum's laboratory (Leiden) and P. Mathis's laboratory (Paris) show a charge-separation complex with a picosecond rise time and lifetimes of around 20–40 ns. The spectrum is a combined P680⁺/pheophytin⁻ spectrum which cannot yet be resolved as the absence of quinones means that no subsequent oxidation of the pheophytin can occur. Reaction-centre triplet formation, first reported by Okamura and Feher¹¹, has been confirmed with material made by Barber's group in my laboratory. The fluorescence-emission spectrum does not show the 695 band which was thought to be diagnostic for photosystem 2. For now it must be presumed that it arises

from light-collecting chlorophyll on CP47.

There seems little doubt that a D1/D2 complex forms the basis of the photosystem 2 reaction centre. The identification is a triumph for comparative molecular biology and must leave biochemists wondering why it took so long, the preparation involving only Triton X-100 treatment and DEAE chromatography. But the photosystem 2 reaction centre is far from solved as the preparation lacks the acceptor quinones. Despite the identification of D1 as the quinone-binding protein, attempts to replace the quinones have not yet succeeded: only preparations retaining CP47 will bind the quinone. The normally tightly bound first electron donor to P680 is lost and has not yet been chemically identified. Cytochrome *b*₅₅₉ is not functional in the preparation and its role in photosystem 2 is unknown, as are the roles of the other poly-

peptides that seem to be essential for assembly of the complex in the thylakoid and the operation of the water-oxidizing enzyme complex. □

1. Namba, O. & Satoh, K. *Proc. natn. Acad. Sci. U.S.A.* **84**, 109–112 (1987).
2. Duysens, L.N.M. *Brookhaven Symp. Biol.* **11**, 10–15 (1959).
3. Reed, D.W. & Clayton, R.K. *Biochem. biophys. Res. Commun.* **30**, 471–476 (1968).
4. Diesenhofer, J., Epp, O., Miki, K., Huber, R. & Michel, H. *J. molec. Biol.* **100**, 385–398 (1984).
5. Diesenhofer, J., Epp, O., Miki, K., Huber, R. & Michel, H. *Nature* **318**, 618–624 (1985).
6. Metz, J. & Bishop, N. *Biochem. biophys. Res. Commun.* **94**, 560–566 (1980).
7. Nakatani, H.Y., Ke, B., Dolan, E. & Arntzen, C.J. *Biochim. biophys. Acta* **765**, 347–352 (1984).
8. Vermaas, W.F.J. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 9474–9477 (1986).
9. Trebst, A.Z. *Naturforschung* **41c**, 240–245 (1985).
10. Satoh, K. & Katoh, S. *Proc. VII Int. Congr. Photosynthesis* (in the press).
11. Okamura, M.Y., Satoh, K., Isaacson, R.A. & Feher, G. *Proc. VII Int. Congr. Photosynthesis* (in the press).

M.C.W. Evans is a professor in the Department of Botany and Microbiology, University College London, London WC1E 6BT, UK.

Plasma physics

A triple-soliton accelerator

T. Tajima

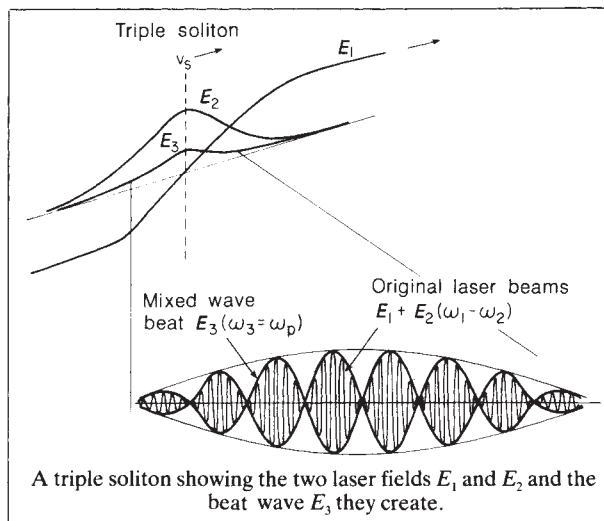
CONVENTIONAL particle accelerators in high-energy physics give a maximum accelerating field — about 1 MeV cm⁻¹ — limited by their technology. Recent calculations show that by using novel effects in laser beams, fields 100–1,000 times larger than this could be achieved.

The mixing of two laser beams of frequencies ω_1 and ω_2 injected into a plasma is known to create a strong plasma-wave accelerating field of frequency $\omega_3 = \omega_1 - \omega_2$. This frequency should be equal to the plasma's own vibrational frequency ω_p to resonate with the plasma wave and enhance it to large amplitudes. Because the plasma wave ω_3 is created by beating between two laser beams, this mechanism for strong accelerating fields is called beat-wave acceleration, and holds promise for the production of very high energies. But

it turns out that because ω_3 has a much slower group velocity v_3 (almost zero) than the incident waves, the main energy of the accelerating structure is left out of the leading laser pulse, creating a turbulent plasma wake behind the laser pulse and so wasting incident laser beam energy. McKinstrie and Dubois (*Phys. Rev. Lett.* **57**, 2022; 1986) have now obtained exact solitary-wave (triple-soliton) solutions of the three beat-wave envelope equations: a triple soliton is a combination of two ordinary transverse electromagnetic waves, and an electrostatic wavepacket on which the charged particles can 'surf' along. Their analytical solution suggests that at an optimal frequency shift there is a complete reabsorption of the plasma wake back into the laser wave. The solitary pulse can curtail the wake of the plasma

wave and improve the efficiency of acceleration of particles.

The idea of a triple soliton is that instead of leaving the energy to flow one way from the two lasers to the plasma wave (and wake) we try deliberately to open up a channel of energy flow from the plasma wave back to the laser beams after appropriate use by the plasma wave. McKinstrie and Dubois classify triple solitons into two categories, temporal and spatial, depending on the velocity of the soliton,



v_s , compared with c , the velocity of light. The spatial soliton has velocity less than c and can be Lorentz-transformed to a co-ordinate frame in which it is solely a function of position and is typically determined by initial-value conditions. The temporal soliton is faster than c and can be transformed to a frame where it is solely a function of time and is typically specified by boundary conditions. McKinstrie and Dubois call triple solitons types A, B and C, depending on the soliton velocity, v_s , relative to the group velocities of the two laser beams, v_1 and v_2 , and the plasma wave, v_3 . They recommend a type B ($c > v_1 > v_2 > v_s > v_3$) spatial soliton for beat-wave acceleration. On the other hand, we (Mima, K. *et al.* *Phys. Rev. Lett.* **57**, 1421; 1986) have previously suggested a triple soliton for beat-wave acceleration but specified type A ($v_s > v_1 > v_2 > v_3$) with v_s matched to c .

For the particles not to outrun the soliton before significant acceleration occurs, v_s has to be close to c . McKinstrie and Dubois also imply that the temporal soliton ($v_s > c$) is not physically relevant because it carries no information: the only usable solitons for acceleration are types A and B spatial solitons. They further infer that because the type A spatial soliton runs into a stiff condition $c > v_1 > v_2 > v_s > v_3$ and $c - v_1 \ll c$, the soliton becomes very short and the envelope approximation may break down. This is why they prefer the type B spatial soliton.

There are several points worth discussing about this choice. The first is the physical realizability of solitons faster than the speed of light; the second is the relative merits of type A and B solitons; and the third is the stability of solitons. A triple soliton propagates with speed v_s that transforms transverse electromagnetic energy into the longitudinal electrostatic wave and back again to the transverse state without transmitting energy at that speed. It propagates by using the energy invested ahead of it. Therefore, surprisingly, a soliton with $v_s \geq c$ does not violate the special theory of relativity. A similar situation arises in a series of linear-accelerator klystrons that are charged up and then fired in sequence so that the propagating electromagnetic wave has any desired phase velocity. Temporal or spatial solitons can be made by choosing appropriate boundary or initial conditions.

According to the theory of nonlinear wave interaction, if the parent wave has a higher frequency ω_1 than the daughter waves ω_2 and ω_3 , energy exchange from parent to daughter can be complete (the decay-instability), whereas if ω_1 is less than ω_2 and ω_3 only a partial exchange of energy from parent to daughters takes place (simple frequency mixing). Now because in type A, ω_1 is the parent ($\omega_1 > \omega_2 > \omega_3$), and in type B, ω_2 is the parent, the energy-exchange process should be more

spontaneous in type A.

An exciting new avenue is opening up in accelerator physics as a result of this work. Plasmas are often too unstable to sustain a stable structure with intense fields. But potentially much more efficient and coherent particle acceleration may be achieved if the self-binding properties of triple solitons and their self-induced transparency can be exploited; two laser beams and a plasma wave, appropriately tailored, can reinforce each other. Such investigation may have applications

in optical fibre soliton research in telecommunications, where the idea is to send a pulse of photons without disintegration over long distances. The new phenomenon of nonlinear accelerating structures could be used for accelerators in the ionosphere; this self-binding structure may propagate stably and efficiently in the ionosphere plasma, where external control is difficult. □

T. Tajima is in the Department of Physics, The University of Texas, Austin, Texas 78712, USA.

Animal behaviour

Leatherback turtle off scale

N. Mrosovsky

MURMURS of surprise at a recent meeting* greeted the announcement by Scott Eckert (University of Georgia) that a leatherback sea turtle had dived to a depth of 1,200 m. In fact the recording device, a combination of a tiny light on a pressure-sensitive arm and a rotating film, had been pushed off the scale. Calibration allowed the 1,000 m mark to be specified accurately and by extrapolating from the relationship between depth and duration of other dives, Eckert and his colleagues estimated that the turtle had dived to a depth of 1,200 m. This is deeper than the 1,140 m for a sperm whale and is probably a record



for an air-breathing vertebrate¹. The same turtle went deeper than 1,000 m on two occasions. In earlier studies Eckert *et al.* had found¹ maximum depths of 475 m. The same pressure-sensitive apparatus, used successfully by G.L. Kooyman (Scripps Inst. of Oceanography) on various animals besides turtles, will now require modification for leatherbacks.

The leatherback turtle, *Dermochelys coriacea*, is a remarkable reptile in other ways. It is the largest living turtle, with maximum weight exceeding 600 kg, and is also the most widely distributed reptile in the world. In the summer it routinely penetrates temperate or cooler waters. It has been reported from Newfoundland to Argentina, from Norway to South Africa, from Russia to Tasmania, and is not benumbed but active in such waters. Measurements on an animal captured off

Nova Scotia show² that leatherbacks are capable of maintaining a body temperature of at least 18°C above an ambient level of 7.5°C. It remains to be determined whether they are warm blooded mainly on account of their size and thermal inertia, or whether they actively regulate their temperature, as suggested by the existence of countercurrent heat exchangers at the junction of flippers and body³.

This thermal physiology is expressed both in the migration of the turtle from the tropics to far northern or southern latitudes and in its deep dives off its nesting beaches. At a depth of 500 m off the Virgin Islands, where Eckert *et al.*¹ did their work, the temperature is 12–15°C, not so very different from the surface temperature in the summer in the northern parts of its range. The reason for the leatherback's excursions, be they far or deep, may also be similar — to feed. This huge animal subsists largely on jellyfish, which are patchily and unpredictably distributed. Whether the leatherback is also feeding when it dives remains to be discovered.

Medusivory raises another question about this already remarkable animal: how does it withstand the toxins it ingests? But the leatherback's feeding habits are not without problems; they frequently mistake floating plastic bags for jellyfish. Sometimes the gut becomes blocked and the turtle dies in an emaciated condition. Elsewhere I have estimated⁴ from data on stomach contents that today 44 per cent of adult leatherbacks have some plastic in their guts. A reptile that can keep itself warm in cold water, travel for thousands of kilometres, dive to more than a thousand metres and gulp down a Portuguese man-of-war, does not necessarily survive eating an inert plastic bag. □

1. Eckert, S. A. *et al.* *Herpetologica* **42**, 381–388 (1986).

2. Frair, W. *et al.* *Science* **177**, 791–793 (1972).

3. Greer, A. E. *et al.* *Nature* **244**, 181 (1973).

4. Mrosovsky, N. *Mar. Turtle Newsletter* **17**, 5–7 (1981).

N. Mrosovsky is in the Departments of Zoology, Psychology and Physiology at the University of Toronto, Toronto, Ontario, Canada M5S 1A1.

*Seventh annual workshop *Sea Turtle Biology and Conservation* 25–27 February 1987, Wekiwa Springs, Florida.

The pulsar in type II supernova 1987A

SIR—We still do not understand type II supernovae. The evolution of massive stars until the central degenerate cores approach the Chandrasekhar mass is by now a carefully studied problem^{1,2} as is the subsequent collapse and conversion of the core to neutrons, releasing energy of the order of 10^{53} erg in thermal neutrinos from the binding energy of the forming neutron star²⁻⁵. This much of the model for type II supernovae has been dramatically confirmed by recent detections⁶ of neutrinos liberated from the core of SN1987A with the predicted⁷ luminosity, temperature and time duration.

But two major mysteries remain. What forces have halted the collapse, driving a shock out through the infalling envelope with sufficient vigour to reverse the motion and accelerate the envelope to the observed speed of 10^4 km s⁻¹? For over two decades a mechanism has been sought. As the original suggestion of Colgate and White⁸ failed, that is, that the centrally emitted neutrinos would couple to outer layers of the core providing enough momentum to reverse the collapse, variants of the original idea continue to be suggested^{9,10}, but there is as yet no clear consensus that any of the recent proposals can prevent collapse of the entire star to a black hole.

Second, what provides the continuous energy input required to maintain the optical luminosity at a very high level for weeks after the initial implosion in the face of adiabatic losses suffered by gas? In SN1987A, expanding from $10^{12.2}$ cm to $10^{15.2}$ cm, the gas thermal energy would be reduced by a factor of 10^6 . For type I supernovae radioactive energy release from ^{56}Ni is thought to provide the delayed energy input, but what is the 'storage battery' in type II events, where radioactivity is not thought to be an important energy input?

The most viable proposed solution^{11,12} uses the specific heat of the vacuum radiation field in an ingenious way. The progenitor star is large (with radius, R , $R > 10^{13.5}$ cm) and may be assumed to have a large additional circumstellar envelope, due to prior mass loss, which reaches an optical depth of unity at a radius of approximately $10^{14.5}$ – 10^{15} cm. Optical brightening occurs when the supernova shock reaches this radius with subsequent energy released from the radiation field stored as $\int aT^4 dV$ below that layer. This solution is perhaps correct for some type II supernovae, but is not possible for SN1987A because the delay between the neutrino burst and the optical brightening of less than 4 h corresponds to a size of $\leq 10^{13.5}$ cm. In any case such 'normal' models always produce a declining

bolometric luminosity whereas SN1987A is increasingly luminous.

With respect to these two major problems our physical understanding of type II events has made only moderate progress since the early 1970s when Ostriker and Gunn¹³ (see also Rees and others¹⁴⁻¹⁶) asked "Do Pulsars Make Supernovae?" proposing that a central pulsar might solve both problems in a straightforward way. The proposal, made more concrete by subsequent work¹⁵⁻¹⁷ remains viable for some type II supernovae, and predictions based on the model can be directly tested by observations of SN1987A. One defect of the 'strong' pulsar model is that the pulsar momentum input is too slow to explain the early large velocity; the rate of acceleration is too slow¹⁷. The early pulsar energy emission rate in the calculations¹³⁻¹⁷ made to date assumed essentially vacuum radiation. This may lead to a severe underestimate of the momentum output if, as proposed by Kulsrud (ref. 19 and personal communication), the appropriate velocity in the radiation formulae is closer to the local Alfvén velocity than the velocity of light.

It remains true that the two shortest period pulsars (with the exception of 'recycled' millisecond pulsars) and youngest pulsars are in young supernova remnants. The Crab nebula is accelerating, with the energy to pump it up coming from the slowdown of the central pulsar. The initial, extrapolated period of the Crab pulsar was 17 ms, about half of its present period. But if neutron stars are typically born rapidly rotating with the angular velocity set by instability to gravitational radiation and viscous losses from non-radial modes, then the initial period would have been a factor of 10 shorter, in the range 0.7–2 milliseconds^{20,21}. The corresponding energy availability would have been 10^2 larger and the nebular expansion velocity a factor of 10 larger than the 1,000 km s⁻¹ for the Crab nebula, comparable to the 13,000 km s⁻¹ velocity observed in SN1987A.

There are two relevant timescales (t_{diff} , t_{ex}), the photon diffusion time through the envelope and the expansion time. These may be written as

$$t_{\text{diff}} = 0.056 \left(\frac{M}{M_{\odot}} \right) \left(\frac{\sigma}{\sigma_t} \right) \frac{\epsilon}{R_{15}} \text{ yr} \quad (1)$$

$$t_{\text{ex}} = 0.32 R_{15} V_8^{-1} \text{ yr}$$

where M is the envelope mass, R_{15} its radius in units of 10^{15} cm, V_8 its velocity in units of 10^8 km s⁻¹ and σ the effective scattering coefficient of the envelope material. Here σ_t is the Thomson scattering cross-section and a shell geometry has been assumed with the envelope thickness $\Delta R = \epsilon R$. From equation (1) it is clear that initially the expansion timescale is shorter,

so that energy injected below cannot be seen. Most of the pulsar energy input is lost to adiabatic expansion. The initial shock has heated the star, and with expansion the thermal content of the star decreases as it is transformed to bulk kinetic energy. The photosphere recedes into the expanding star through successive layers, exposing fresh material. Luminosity decreases due to the adiabatic losses and colours redden due to the increasing photospheric radius and decreasing luminosity.

The two timescales become equal at

$$R_{\text{eq},15} = 0.42 \left(\frac{M}{M_{\odot}} \right)^{1/2} \left(\frac{\sigma}{\sigma_t} \right)^{1/2} \times V_8^{1/2} \epsilon^{1/2} \quad (2)$$

$$t_{\text{eq}} = 0.13 \epsilon^{1/2} V_8^{-1/2} \times (M/M_{\odot})^{1/2} (\sigma/\sigma_t)^{1/2} \text{ yr} \quad (3)$$

at which time the optical depth is still large; $\tau_{\text{eq}} = 300 V_8^{-1} \epsilon^{-1}$. As representative numbers for SN1987A take a bulk velocity of 13,000 km s⁻¹ and a mass of $25 M_{\odot}$ to give (with $\epsilon \sim 1/3$); $R_{\text{eq},15} = 4.4$; $t_{\text{eq}} = 0.10$ yr = 37 days, $\tau_{\text{eq}} = 70$. After that time (the beginning of April for SN1987A) the surface optical luminosity should begin to increase, as one sees the pulsar input more directly. Eventually the envelope acceleration will be detected and after a time

$$t_{\text{tr}} = 2.31 \left(\frac{M}{M_{\odot}} \right)^{1/2} \left(\frac{\sigma}{\sigma_t} \right)^{1/2} V_8^{-1} \text{ yr} \approx 3.2 \left(\frac{\sigma}{\sigma_t} \right)^{1/2} \text{ yr} \quad (4)$$

the envelope will be sufficiently transparent to reveal the extremely hot non-thermal nebula below. In recent work McCray *et al.*²² have described how such an object would be bright in emission lines emitted by the envelope material excited by ultraviolet from below.

The luminosity at maximum is expected in this picture to be several magnitudes brighter after t_{eq} than before, the exact value depending on the pulsar luminosity. The predictions of Ostriker and Gunn¹³, which for simplicity we will not alter to fit the case at hand, indicate for their 'standard case' an initial pulsar angular velocity of 3,000 rad s⁻¹ or period P , of 2 ms, and a corresponding luminosity of $L_{\text{SN}} \approx 10^{10.1} L_{\odot} (P/2 \text{ ms})^{-4}$. If these numbers are correct, we may expect SN1987A to brighten to $m_{\text{bol}} \leq -0.5$ in the near future. This brightening may be characteristic of type II supernovae. It may have been observed in some cases and perhaps missed in others due to the inefficiency of our detection methods for supernovae for magnitudes below maximum light.

Currently^{23,24} there is some evidence to

suggest an overall optical brightening of SN1987A, a property which is quite difficult to understand in models where the energy input is impulsive. Radioactive energy input (W. D. Arnett, preprint, 1987) will retard the decline and may even produce a slight increase in luminosity, so the light curve to date may not yet require pulsar input.

But if there is an embedded rapidly spinning pulsar, then by August 1987 SN1987A should have become several magnitudes more luminous than it was in March and should ultimately show a spectrum with an extremely hard and polarized continuum from the underlying nebula. The nebular magnetic field will be quite large, in the range of 10^0 – 10^2 gauss. Soon after that a very short period pulsar should emerge. As radio luminosity is normally small and not a strong function of period, and the nebular plasma frequency will be large, a radio pulsar will not be detected²⁵. In the case of the Crab nebula a significant fraction of the slowdown energy is seen as hard pulsed radiation, so we may expect X-ray and perhaps optical pulses in the millisecond range to be seen from SN1987A²⁵ by 1990.

The relatively specific predictions presented here (submitted 3 April 1987) allow one to test the hypothesis^{12–16} that the SN1987A supernova is being driven by a central pulsar.

I thank D. Arnett, J. N. Bahcall, R. Durisen, J. E. Gunn, R. McCray and B. Paczynski for useful discussions, and NASA and NSF for material support.

J. P. OSTRICKER

Princeton University Observatory,
Peyton Hall,
Princeton,
New Jersey 08544, USA

1. Paczynski, B. *IAU Symp.* No. 66 (eds Taylor, R. J. & Hessler, J. E.) 62 (Reidel, Dordrecht, 1974).
2. Arnett, W. D. *IAU Symp.* No. 66 (eds Taylor, R. J. & Hessler, J. E.) 150 (Reidel, Dordrecht, 1974).
3. Bethe, H. A. & Wilson, J. R. *Astrophys. J.* **295**, 14 (1985).
4. Burrows, A. & Lattimer, J. M. *Astrophys. J.* **307**, 178 (1986).
5. Woosely, S. F., Wilson, J. R. & Maygle, R. *Astrophys. J.* **302**, 19 (1986).
6. The Kamiokande-II Collaboration *Phys. Rev. Lett.* **58**, 140 (1987).
7. Bahcall, J. N., Dar, A. & Piran, T. *Nature* **326**, 135 (1987).
8. Colgate, S. A. & White, R. H. *Astrophys. J.* **143**, 626 (1966).
9. Bethe, H. A. & Wilson, J. R. *Astrophys. J.* **295**, 14 (1985).
10. Goodman, J., Dar, A. & Nussinov, S. *Astrophys. J.* (in the press).
11. Falk, S. W. *Astrophys. J.* **226**, L133 (1978).
12. Chevalier, R. *Astrophys. J.* **251**, 159 (1981).
13. Ostriker, J. P. & Gunn, J. E. *Astrophys. J.* **184**, L95 (1971).
14. Rees, M. J. *Astrophys. J.* **6**, L55 (1970).
15. Bodenheimer, P. & Ostriker, J. P. *Astrophys. J.* **191**, 465 (1974).
16. Bisnovati-Kogan, G. S. *IAU Symp.* No. 66 152 (eds Taylor, R. J. & Hessler, J. E.) 152 (Reidel, Dordrecht, 1974).
17. Gaffet, B. *Astrophys. J.* **216**, 852 (1977).
18. Symbolist, T. M. D. *Astrophys. J.* **285**, 729 (1984).
19. Kulsrud, R. *Astrophys. J.* **163**, 567 (1971).
20. Friedman, J. L., Ispier, J. R. & Parker, L. *Astrophys. J.* **304**, 115 (1986).
21. Cutter, C. & Lindblom, L. *Astrophys. J.* **314**, 234 (1987).
22. McCray, R., Shull, J. J. & Sutherland, P. *Astrophys. J.* (in the press).
23. *IAU Circ.* No. 4352 (1987).
24. *IAU Circ.* No. 4356 (1987).
25. Bahcall, J. N., Rees, M. J. & Salpeter, E. L. *Astrophys. J.* **162**, 737 (1970).

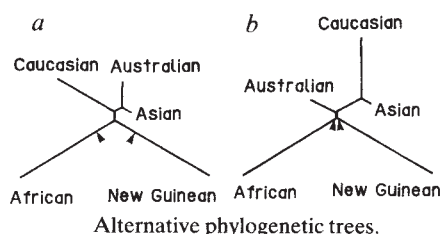
Time and place of human origins from mt DNA data

SIR—Cann *et al.*¹ claim, from their mitochondrial DNA data, that the origin of current human populations was probably in Africa and that the time of divergence was about 200,000 years ago. We think however, that the place and the time of origin of *Homo sapiens sapiens* are difficult to estimate from their data.

The figure shows two phylogenetic trees obtained from the genetic distance data of Table 1 of ref. 1. These two unrooted trees were constructed by the neighbour-joining method² and both are equally the best of 15 possible topologies. Branch lengths are proportional to the estimated distances. Note that Africans and New Guineans cluster together in both trees. Although the branch separating these two populations from others is short in both trees, this clustering is supported by other tree-making methods such as UPGMA, Fitch and Margoliash's method, Farris' distance Wagner method, modified Farris method, and the transformed distance method (see ref. 3 for details of these methods). However, analyses of data from many nuclear loci show that New Guineans are clustered with Australians (Aborigines), instead of Africans^{4,5}.

Arrows in the figure denote the points of the root, the common ancestor for all populations, determined by minimizing the variance of the distances from all populations. In this case, a rough constancy of the rate of evolution is assumed. Interestingly, there are two such points in both trees of the figure. If we accept this rooting procedure, not only Africans but also New Guineans can be the population that diverged first.

Without referring to the population phylogeny described above, Cann *et al.* focused their discussion on the gene phylogeny of mitochondrial DNA types (Fig. 3 of ref. 1). To give a tentative time scale, the rate of nucleotide substitution was assumed to be 2–4% per nucleotide site per million years. However, if we apply this rate to the data on hominoids⁶, the divergence time between humans and chimpanzees or gorillas becomes only 1.4–2.8 million years. A more reasonable estimate (0.71% per site per million years) has been obtained by Nei⁵. If we take this value, the divergence time of the point 'a' of Cann *et al.*'s Fig. 3 (ref. 1) becomes 400,000 years. This is in the middle of the supposed time span of *Homo erectus*, but



this may not mean that this alternative estimate of the rate is too small. Genetic polymorphism is expected in the ancestral human population and the mean coalescence time of alleles (mitochondrial DNA types) in this ancestral population can exceed one million years, if the average effective size and the generation time in the past have been 10,000 and 25 years, respectively⁷. Thus, there is no need to assume multiple colonizations.

In conclusion, we believe that data from mitochondrial DNA, which can be regarded as a single locus, are not enough by themselves to establish the branching pattern of human populations.

NARUYA SAITOU
KEIICHI OMOTO

Department of Anthropology,
Faculty of Science,
The University of Tokyo, Hongo,
Bunyo-ku, Tokyo 113, Japan

1. Cann, R. L. *et al.* *Nature* **325**, 31–36 (1987).
2. Saitou, N. & Nei, M. *Molec. Biol. Evol.* (in the press).
3. Nei, M. *Molecular Evolutionary Genetics* (Columbia University Press, New York, 1987).
4. Omoto, K. in *Molecular Evolution, Protein Polymorphism and the Neutral Theory* (ed. Kimura, M.) 193–214 (Springer, Berlin, 1982).
5. Nei, M. in *Population Genetics and Molecular Evolution* (eds Ohta, T. & Aoki, K.) 41–64 (Japan Science Society Press, Tokyo, 1985).
6. Ferris, S. D. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **78**, 2432–2436 (1981).
7. Takahata, N. & Nei, M. *Genetics* **110**, 325–344 (1985).

Varieties of sexual contact

SIR—May and Anderson report a wide variation in the estimates of β , the transmission probability for a single contact between a carrier of human immunodeficiency virus 1 (HIV-1) and a susceptible person¹.

One explanation may lie in the failure of researchers, including May and Anderson, to distinguish between the various degrees of contact. In other words, it is not sufficient to ask how often new partners are acquired. One should rather inquire how often a new 'risk-partner' is acquired². Moreover, a new risk-partner is anyone with whom one has had unsafe sex after having had unsafe sex with anyone else. This applies (perhaps repeatedly) even to one's regular partner in an otherwise exclusive relationship.

The problem is highlighted by the fact that male homosexual practices vary with different partners³. Because, for example, casual gay sex does not usually involve penetration, whereas casual heterosexual sex does⁴, any estimation of β which ignores this point may easily be in error by an order of magnitude.

GORDON REECE

Department of Engineering Mathematics,
University of Bristol,
Bristol BS8 1TR, UK

1. May, R. M. & Anderson, R. M. *Nature*, **326**, 137–142 (1987).
2. Reece, G. J. *Modelling the Spread of AIDS*, (Social Research Unit, University College, Cardiff, 1987).
3. McManus, T. J. & McEvoy, M. *Br. J. sex. Med.* **14**(4), 110 (1987).
4. Kinsey, A. C., Pomeroy, W. B. & Martin, C. E. *Sexual Behaviour in the Human Male* (Saunders, Baltimore, 1948).

NASA: decline and fall

John Noble Wilford

Challenger: A Major Malfunction. By Malcolm McConnell. Doubleday, New York/Simon & Schuster, London:1987. Pp.269. \$17.95, £12.95.

Prescription for Disaster: From the Glory of Apollo to the Betrayal of the Shuttle. By Joseph and Susan Trento. Crown, New York:1987. Pp.320. \$18.95.

THE Rogers Commission, concluding its thorough investigation, called the Challenger disaster "an accident rooted in history". The investigators were referring primarily to the "flawed" management of the space shuttle programme itself and the persistent neglect of manifest deficiencies in the booster O-rings, which turned out to be the direct culprits. Judging by the many post-accident revelations, however, these were merely symptoms of a deeper and more debilitating condition besetting the US civilian space programme. It is this history, of the decline and fall of the once-proud NASA, that is the subject of these two books by American journalists.

These are the first books to deal with the explosion on 28 January 1986 which destroyed the Challenger, killed its crew of seven and left the entire space programme in shambles. Not surprisingly, both accounts suffer from the rush to publication. They are rich in the disturbing and depressing detail of what had gone wrong with the shuttle project and with NASA in general, but for the most part superficial in their analysis of conditions before and after the accident. Even so, the authors are to be commended for managing in a short time to gather so much evidence to document the sorry state of affairs within NASA in the years and months before that cold morning in January.

The books complement each other. *Challenger: A Major Malfunction* retells the story of the events leading up to the decision to launch the space shuttle over the objections of some engineers. The emphasis is on the immediate causes of the accident and the people involved in the tragic decisions. As the title promises, *Prescription for Disaster: From the Glory of Apollo to the Betrayal of the Shuttle* pieces together more of the early history of NASA's operations, when political support was strong and quality control the passion, and then gives a relentless accounting of how this tradition was betrayed by a succession of political retreats and technological compromises. Against this background, the events immediately preceding the accident unfold with a force of inevitability.

Of the two, *Challenger: A Major Malfunction* is the more successful book. Malcolm McConnell, who has reported on space exploration for *Reader's Digest*,

writes with a crisp style and a sense of drama. He organizes the narrative to maintain suspense even when the outcome, and many of the details, are already well known. His reflections on the causes and effects of the accident are few but graphic. "Here is policy made tangible", McConnell writes on seeing the Challenger wreckage spread out on the floor of

and, it seems, even deceiving itself. Meanwhile, the agency's leaders were pre-occupied with selling the White House and Congress yet another manned-flight venture, the space station, and left the running of the shuttle to the lower echelons within NASA.

The lower echelons suffered from the bitter rivalries among NASA's field centres. In one of the most illuminating chapters, McConnell describes the dictatorial leadership of William R. Lucas, director of the Marshall Space Flight Center in Alabama. One of his rules, which came to light after the accident, was that, under no circumstances, was the Marshall Center to be the cause of delaying a launch (p. 109). This probably explains the attitude of Lawrence Mulloy, a Marshall manager, the night before the Chal-

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Wreckage from Challenger — "Here among the debris... one is forced to recognize that political compromise, bureaucratic deception, and corporate duplicity all have consequences."

a hangar at the Kennedy Space Center in Florida (p. 252). "Here among the debris of the Orbiter one is forced to recognize that political compromise, bureaucratic deception, and corporate duplicity all have consequences."

The substance of the book bears out this assessment. The author recalls how in the early 1970s the space agency, desperate to assure its survival by winning support for a new manned-flight project, oversold the shuttle as a reusable space transport system that would revolutionize the economics of space travel. President Nixon gave his reluctant approval, but his and subsequent administrations provided too little financial support for such a great undertaking. NASA was thus forced to cut corners in design and testing, and to sacrifice much of the internal quality-control structure that had contributed to its successes in the 1960s. Then, as it became apparent that the shuttles were falling woefully short of economic and operational expectations and might be taken over by the military, NASA's management pushed even harder to maintain unrealistic schedules, deceiving the public

lenger launch. Losing patience with contractor engineers who expressed concern about O-ring performance in cold weather, Mulloy finally blurted out: "My God, Thiokol, when do you want me to launch? Next April?"

McConnell falters in the end when he writes, "After the accident, a new, saner space policy was emerging". Perhaps he wanted desperately to conclude on an optimistic note. True, the failed policy of dependence on a sole launching system, the shuttle, has been abandoned. But so far little else seems to have changed. NASA continues to be an agency adrift, its independence threatened by the growing Pentagon space programme and its future clouded by a lack of direction and real support from the Reagan administration. In its struggle to win backing for the space station, NASA has once again been caught in a gross underestimate of projected costs.

Prescription for Disaster is poorly written, a rambling and largely undigested work, but a rich mine of raw material for future histories of the US space programme. Joseph and Susan Trento

interviewed all the former NASA administrators as well as several dozen veterans who worked at the agency over the past 30 years. Their stories read like an oral history of the petty disputes, bureaucratic turf battles and other dirty linen that seemed especially to mark NASA's shuttle era.

One example should be sufficient. With all the problems of the shuttle on his mind, James M. Beggs, NASA administrator from 1981 to 1985, had to suffer repeated interference from the White House over such trifling matters as VIP guest lists for launchings and the "cola wars". In his interview with the Trentos, Beggs told of the "aggravation" caused by Michael Deaver, the lobbyist and close friend of the Reagans, who sought to have his client Coca-Cola favoured over Pepsi-Cola in a zero-g test of soft-drink cans (pp.256–257). One feels sympathy for Beggs, but is left wondering why NASA got itself into such a silly predicament in the first place.

The clear impression here is that NASA sold its soul in return for the shuttles. According to the Trentos, the agency set high standards for technical excellence and became the symbol of boldness and success because of its early leaders, people such as T. Keith Glennan, James E. Webb, Robert Seamans, Rocco Petrone and Robert R. Gilruth. Of Gilruth, former director of the Houston space

centre, the authors write that he was the "living embodiment of what NASA was supposed to be — effective, curious, and humane" (p. 75).

The Trentos may have romanticized the early NASA to stress its more recent decline. They can find no heroes of the shuttle era. The cumulative effect of the stories told by these interviewees is one of indecision, compromise, political meddling and a tragic neglect of engineering standards. A deplorable trend noted by the authors is the frequency with which agency officials move to and from employment with industry contractors, raising questions of potential (perhaps real) conflicts of interest.

A reader can draw two lessons from these books. One lesson — and this strikes home to this reviewer — is that we must not be so dazzled by the spectacle of space flight as to accept at face value everything NASA, or any agency, tells us. Nearly everyone, from the White House to Congress to the news media, fell into this trap. Consequently, NASA could delude itself into believing it was well-nigh infallible.

The other lesson, which has yet to sink in, is that if spacefaring is a worthy endeavour it is worth doing well. □

John Noble Wilford, an author and journalist, is in the Science News Department, New York Times, 229 West 43rd Street, New York, New York 10036, USA.

Pattern and rhythm

John J. Tyson

Molecules, Dynamics & Life: An Introduction to Self-Organization of Matter. By A. Babloyantz. Wiley: 1986. Pp.345. £50.45, \$55.

SINCE the publication of a short report in *Nature* in 1953, the attention of biologists has been drawn increasingly to macromolecular structure and function — rightly so, considering the success of this approach and the mind-stretching opportunities opened up by genetic engineering. But the structure/function paradigm of modern molecular biology provides a basically static view of life. Daily cycles of sleep and wakefulness, precise temporal and spatial organization in developing embryos, steady waves of contraction that sweep across a beating heart, such commonplace rhythms and patterns of living organisms are largely missing from the conception of life as structure, code, assembly. Over the past 20 years a study of the dynamics of biological systems, of their spatio-temporal organization, has taken root and spread, particularly in continental Europe. In *Molecules, Dynamics & Life* Agnes Babloyantz tells the story of this promising approach to the science

of life in terms intelligible to the non-specialist: the chemist, engineer or biologist who has heard rumours of something unusual and would like to know more about it.

Her job is not an easy one because the study of biological dynamics naturally expresses itself in terms of nonlinear differential equations, bifurcation theory, strange attractors and other arcane furnishings of dynamical systems theory. To make matters worse, she first plunges into the recondite subject of nonequilibrium thermodynamics. However, Dr Babloyantz proves herself to be a pleasant, practical and reliable guide to new territory which is still largely uncharted and inhospitable to tourists. Her style falls halfway between that found in a popular account and that of a textbook. She tells her story in a chatty, down-to-earth way, while also giving serious scientific consideration to fundamental issues of the self-organization of matter.

The first fundamental issue is the thermodynamical problem of how spatial and temporal order can arise spontaneously in physicochemical systems obeying the second law of thermodynamics (which requires closed, isolated systems to proceed monotonously to an equilibrium state of maximum disorder). The simple answer to this question, which was already apparent in the last century, is

that spontaneously organizing systems (such as living cells) are neither closed nor isolated. Following the Brussels school of nonequilibrium thermodynamics, Babloyantz fleshes out this idea, showing in a clear, logical fashion how open systems operating far from equilibrium can lose stability of the "thermodynamic" branch of steady states and evolve spontaneously to regimes of spatial and temporal organization.

As Babloyantz points out, however, nonequilibrium thermodynamics does not provide a practical way of studying dissipative structures appearing beyond the region of stability of the thermodynamic branch. The work-a-day tools for such studies are provided by chemical kinetics, the qualitative theory of differential equations and numerical analysis. So she explains how to use these tools and then employs them on a variety of sample problems: oscillations and pattern formation in the Belousov-Zhabotinskii reaction, the primordial evolution of biological macromolecules, glycolytic oscillations in yeast cell preparations, aggregation of slime mould amoebae, morphogenetic models of positional information in embryos and a neural network model for epileptic seizures.

Dr Babloyantz has produced an engaging and earnest introduction to the field of self-organization in chemical and biological systems. Because she assumes that her readers will have no prior knowledge of biology but some familiarity with differential calculus, and because she concentrates in Part I on thermodynamic issues, her introduction is most suitable for physical scientists. But biologists could also benefit from her guidance if they skip straight to Part II and are willing to smile indulgently at some simplistic explanations intended for a different reader. □

John J. Tyson is a Professor in the Department of Biology, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061, USA.

New journals review

On 24 September *Nature* will publish the seventh annual review supplement devoted to science journals.

Criteria for inclusion of a journal in the 1987 issue are that:

- (i) the first number appeared, or the journal was retitled, between June 1985 and May 1986 (the second cut-off date allows at least three issues of a journal to have been published, the minimum number on which a reasonable judgement can be based);
- (ii) it is published at least three times a year;
- (iii) the main language used is English.

Publishers and learned societies are invited to send four different issues of each suitable periodical, including the first and most recent numbers (if from outside the United Kingdom, by air mail) to: The Review Editor, *Nature*, 4 Little Essex Street, London WC2R 3LF, England. Subscription details for 1987 (and 1988 if possible) should be included.

The open question in selenology

David W. Hughes

Origin of the Moon. Edited by W. K. Hartmann, R. J. Phillips and G. J. Taylor. Lunar & Planetary Institute, 3303 NASA Road 1, Houston, Texas 77058:1986. Pp.781. \$25.

EARTH'S satellite Moon has a mass of 0.012453 Earth masses, a diameter of 0.27252 Earth diameters and, at present, is about 384,401 km away, this distance increasing by 4 cm per year. It has been assiduously observed throughout civilization, culminating in the period between July 1969 and December 1972 when 12 men walked on its surface and about 382 kg of rock was returned for analysis on Earth. In astronomical terms, therefore, the Moon must be classed as a well-known object, but astronomers still have to admit shamefacedly that they have little idea as to where it came from. This is particularly embarrassing, because the solution of the mystery was billed as one of the main goals of the US lunar exploration programme.

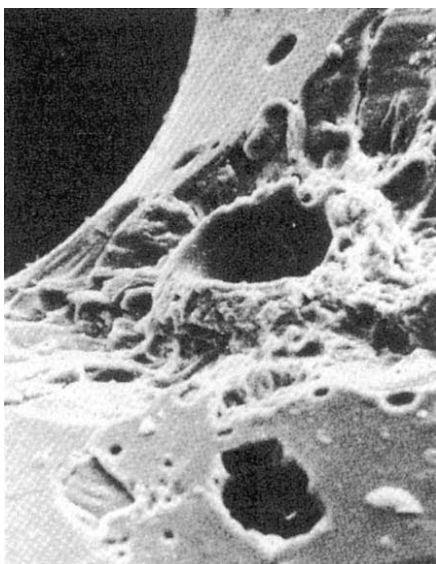
Science is replete with "if only's". In the lunar case things might have been different if only some primordial "genesis" rock, 4.5×10^9 years old, had been found; chemical, isotopic and petrologic analysis of the rock might have provided the necessary clues. Astronauts were trained to look for it, but only a few chips were found. The extraordinarily intense lunar bombardment, which was associated with the sweeping up of planetesimal debris and which occurred between 4.5×10^9 and 4.0×10^9 years ago, blanketed the Moon with a pulverized regolith and hid the evidence.

Twelve years after the last lunar landing was as good a time as any to hold a conference on the subject of selenology. The conference was organized by the Lunar and Planetary Sample Team and took place in Kona, Hawaii, in October 1984. Thirty three papers from the meeting are published in the volume under review, and in one way the reviewer's job is easy — the book cannot be compared with others on the subject because there are none.

As with most studies of origin mechanisms, the starting point must be the known facts. Moving beyond the mass, size and angular momentum, data that can easily be gleaned from Earth observation, a large number of important clues have come from sample analysis. Lunar rocks and soils are depleted in volatile elements, much more so than those on Earth, this loss probably being caused by some highly energetic process that heated lunar material to high temperatures. Iron has been

depleted to a value of only 25 per cent of what one would expect from cosmic abundance. The oxygen isotopic ratio is similar to that of Earth. And comparing lunar material with rocks from the Earth's mantle it has been found that there are strong similarities between the relative abundances of the siderophile elements (such as Co, Ni, W, P, S, Se, Te) and lithophile elements (V, Cr, Mn).

Plagioclase takes up to 75 per cent of lunar surface material and could only have been produced by the partial melting of a crustal layer several hundred kilometres thick followed by crystal fractionation. This again points to a high-energy origin process. If the celestial clock is wound back it is clear that 4.5×10^9 years ago the Moon was very much closer to the Earth — about 10 Earth radii away (60,000 km)



Zap crater, the result of the impact of a high-speed meteoritic dust grain on Moon rock. The picture is reproduced from Discovering the Universe, by William J. Kaufmann III, a textbook for a one-term, descriptive course on astronomy. Publisher is W. H. Freeman, price is pbk \$24.95, £24.95.

— with a direct orbit inclined at 10° to the Earth's equatorial plane. About 400 pages and 16 papers in the book are taken up by a discussion of the dynamical, geochemical, petrologic, isotopic, thermal and magnetic facts relating to the Earth-Moon system. The remainder covers ideas about lunar origin.

Five theories are studied in detail. According to the first of them, "intact capture", the Moon was formed somewhere else in the Solar System and originally had a heliocentric orbit. It was then captured intact by the Earth. Originally it was thought that this theory could account for a very exotic lunar composition. Unfortunately the capture process is highly dependent on velocity, requiring a very low relative velocity between Earth and Moon. This occurs only if the Moon's original orbit had a radius between 0.95 and 1.05 AU — just like that of Earth, so

they would probably have been formed out of the same material anyway.

The second theory, "co-accretion", holds that Earth and Moon formed at the same time, the Moon accreting from a disk or swarm of solid particles that were orbiting the Earth. Formation of the Moon would be like planetary formation, but on a different scale. Unfortunately, in this model there are difficulties in explaining the differences in composition, providing the energy to melt the magma and supplying the angular momentum. Collisions between particles in the disk could, however, have led to devolatilization.

"Earth fission", the third theory, was first proposed by G. H. Darwin in 1879. He thought that free oscillations in a fast-spinning (~ 4 hour) proto-Earth would be influenced by solar tides, and would produce a tidal bulge which grew higher and higher until the tip of a highly elongated proto-Earth broke off to form the Moon. Modern versions invoke rotational instability in a proto-Earth which had a 2.6-hour day. Core segregation caused the spin rate to increase, making the shape distort until the more tenuously connected end broke off. The Moon probably did not break off intact. According to a more likely model, dispersed material was shed from the elongated Earth, this material forming a disk which subsequently condensed to form the Moon. The main problem with this idea is that the Earth-Moon system needs four times more angular momentum than it has at present.

The fourth theory is "collision ejection", which has the Earth being hit by a planetesimal of about 0.1 Earth masses. A large expanding cloud of vapour and dust was produced by this collision. Some fell back to Earth, some escaped into space, but enough of it went into orbit eventually to accrete to form the Moon.

Finally there is "distinguishative capture", in which a loosely bound planetesimal passed close enough to Earth to enter the Roche limit, was tidally disrupted and part of the debris was subsequently captured and condensed to form the Moon.

All of these theories are reviewed in detail in *Origin of the Moon*. This is a fascinating subject, and it has been well served by both authors and editors — the book is well produced, thorough, carefully structured, up to date and well referenced. It shows that scientific fashion wanders from theory to theory, all of which have good and bad points. At the moment co-accretion is retreating from favour and collision ejection is in the ascendant. But astronomy would be rather dull if we knew all the answers and it is clear from this book that "Where did the Moon come from?" is still an open, exciting and challenging question. □

David W. Hughes is Senior Lecturer in Astronomy and Physics, Department of Physics, University of Sheffield, Sheffield S3 7RH, UK.

Missionary work in the rain forest

Norman Myers

The Geophysiography of Amazonia: Vegetation and Climate Interactions. Edited by Robert E. Dickinson. Wiley:1987. Pp. 544. £69.40, \$72.50.

AMAZONIA is such a highly integrated ecosystem, with myriad interactions at multiple levels, that we — whether scientists, conservationists or developers — need to approach it with as much interdisciplinary spirit as we can generate.

"So what else is new?" might be the response of a cynic. Three things are new. First, we are beginning to understand just how little we understand about Amazonia. For sure, we know much more than we did in 1970, notably about the region's soils, its nutrient flows and other components of the system. But we have all too few insights into how the components fit together, how they *work*. Secondly, much of Amazonia is undergoing extremely rapid change. Some areas are suffering deforestation at rates that have increased by a factor of 13 in just 10 years. Thirdly, and arising from the other two, Amazonia is presenting a greater challenge to science than has sometimes been supposed. Its predominant feature is not so much its biotic richness, with its millions of species, but its ecological complexity, with intricate interaction of those species among themselves and with their physical environments. This last factor applies especially to the dynamic linkages between the region's vegetation and its climate. Hitherto it has been widely believed that forests do not affect rainfall. Now evidence is emerging that there could be feedbacks between deforestation and precipitation patterns.

To generate a response of sufficient scope, science needs to develop an intellectual framework that, according to James Lovelock, one of the contributors to the book under review, should rank as "a new discipline of geophysiography". The field should encompass the phenomena, functions and processes of the biosphere and its support systems; and its study should be a mission-orientated affair,

New in paperback

Harvard University Press has re-issued three of its best-selling science books from 1985:

- *Revolution in Science* by I. Bernard Cohen (reviewed in *Nature* 315, 433; 1985). Price is \$9.95, £7.95.
- *Niels Bohr: A Centenary Volume* edited by A.P. French and P.J. Kennedy (reviewed in *Nature* 320, 221; 1986). Price is \$14.95, £11.95.
- *The Dialectical Biologist* by Richard Levins and Richard Lewontin (reviewed in *Nature* 320, 23; 1986). Price is \$8.95, £7.25.

"directed toward procedures for the diagnosis and prevention of planetary maladies". Nothing like being imaginative, not to say ambitious.

To promote this cause, in 1985 the United Nations University convened a conference "Climatic, Biotic and Human Interactions in the Humid Tropics", with emphasis on Amazonia. This book presents the proceedings, and a fine publication it is. Broad ranging in purpose, eclectic in spirit and pioneering in accomplishment, it meets the needs of the times. After all, if rampant deforestation in Amazonia triggers climatic dislocations, whether in the region itself or elsewhere in Brazil (what if there were less rainfall in southern Brazil with its rich farmlands?), the repercussions could be pervasive and irreversible. As several papers in the book show, an outcome of this sort is far from unlikely — and there may even be effects on climate further afield.

In addition to Lovelock's overview paper, the book contains a first-rate account of some climatic mechanisms in the humid tropics by Robert Dickinson,

who also serves as the book's editor; an up-to-date review of deforestation in Amazonia by Philip Fearnside; a critical assessment of the tropics' contribution to atmospheric chemistry by Paul Crutzen; and several illuminating analyses of biogeochemical cycles, soil productivity, plant-animal interactions, micrometeorology, trace gases, albedo, general circulation models and hydrological cycles by luminaries such as Wolfgang Seiler, Rattan Lal, Ghilleen Prance, Luiz Molion, Eneas Salati and Ann Henderson-Sellers.

I wish there had been more on remote-sensing technologies for inventory and monitoring research, on the biotas as a growing source of atmospheric carbon dioxide and on "fertilizer effect" feedbacks of a greenhouse-affected biosphere. But no matter. This is a splendid book which portends well for Wiley's series "Climate and the Biosphere". □

Norman Myers, Upper Meadow, Old Road, Headington, Oxford OX3 8SZ, UK, is a consultant on development and environmental matters.

Take Manhattan

Clive Ellory

The Ca²⁺ Pump of Plasma Membranes. By Alcides F. Rega and Patricio J. Garrahan. CRC Press: 1986. Pp.173. \$76 (United States), \$87 (elsewhere).

ALTHOUGH control of intracellular calcium levels is central to current ideas in cell physiology, the calcium pump of plasma membranes has been reviewed to a far more limited extent than its close relatives, the sodium pump and the calcium pump of the sarcoplasmic reticulum. This monograph by Rega and Garrahan sets out to redress the situation.

Stylistically it is an unusual work. It lacks a preface, introduction or concluding summary, and is presented strictly as a review, although illustrated with carefully chosen, benchmark figures. In places the book is very good. There is a brief and lucid description of the development of ideas which gives the relevant experimental evidence. However, as is so often the case with monographs impinging on vast topics such as the control of intracellular Ca, the problem is unevenness. The authors begin with chapters on the methodology of Ca measurement, and the role of Ca in cellular function. The first of these suffers from obsolescence; there is, for example, no mention of fura-2. The second is an attempt to cover such an enormous area that omissions and imbalance are inevitable. Thus it seems unjustified to describe Ca-activated K-channels almost exclusively in terms of the

red cell Gardos channel, dismissing all other cell types in a few sentences.

In fact it is clear that the authors' first love is the erythrocyte, and that they are happiest with enzyme kinetics and detailed partial reactions of the red-cell Ca pump, which are covered fully and very well. Certainly, the chapter on purification justifies the choice of the erythrocyte for most biochemical studies. It also emphasizes the low membrane abundance of this calcium pump, compared with other transport systems. The importance of calmodulin arises in two contexts, purification and regulation. Finally other modulators and inhibitors are also considered.

In general the writing is clear and readable, although there are some rather untidy orthographic errors and an occasional awkwardness stemming from too-direct translation of Spanish syntax. Overall this is a useful book, which will be a convenient review for those interested in the plasma membrane Ca ATPase.

In conclusion, I must say that the partisan emphasis on the red cell reminded me of the famous *New Yorker* cartoon of the world where Manhattan firmly occupies most of the foreground, the East Coast, the rest of the United States and the world being represented with progressively greater compression. As a self-confessed devotee of the red cell I appreciate the importance of the erythrocyte as a model system. But in the context of calcium transport other cell types should also get their fair share of attention. □

Clive Ellory is Reader in the University Laboratory of Physiology, Parks Road, Oxford OX1 3PT, UK.

Prevalence of *ras* gene mutations in human colorectal cancers

Johannes L. Bos, Eric R. Fearon*, Stanley R. Hamilton†, Matty Verlaan-de Vries, Jacques H. van Boom‡, Alex J. van der Eb & Bert Vogelstein*§

The Department of Medical Biochemistry and ‡ Department of Organic Chemistry, State University of Leiden, Sylvius Laboratories, The Netherlands

* The Oncology Center and † Department of Pathology, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA

A combination of DNA hybridization analyses and tissue sectioning techniques demonstrate that ras gene mutations occur in over a third of human colorectal cancers, that most of the mutations are at codon 12 of the c-Ki-ras gene and that the mutations usually precede the development of malignancy.

ONE of the exciting recent developments in cancer research has been the identification of specific point mutations that occur in human tumour cells. These mutations have been exclusively localized to amino acids 12, 13 and 61 of the three *ras* genes, c-Ha-*ras*, c-Ki-*ras*, or N-*ras*, (reviewed in refs 1–3). Some of these same mutations have been shown to be present in the *ras* genes of RNA tumour viruses and in the cellular *ras* genes of animal tumours induced by carcinogens^{1–3}. Furthermore, mutated *ras* genes have been shown to transform immortalized cell lines and to act in a cooperative fashion with other oncogenes to transform primary cells in culture^{4–6}.

These data raise important questions about the function of *ras* gene mutations in common human tumours. Several studies have reported on the frequency of *ras* mutations in epithelial cancers of humans; only 2 of 38 transitional cell carcinomas of the bladder⁷, none of 20 mammary adenocarcinomas⁸ and none of 26 gastric carcinomas⁹ had activated *ras* genes. These studies have been interpreted as indicating that *ras* genes are activated in only a small fraction of human tumours. An alternative explanation, however, is that activated *ras* genes in some human tumours may have escaped detection in the NIH 3T3 transfection assay which was used in all the studies noted above. In solid tumours, two factors might prevent detection of activated *ras* genes by transfection assays. First, the DNA from primary human tumours is often slightly degraded, both because of necrosis *in vivo* and degradation occurring during surgery and subsequent tissue handling. For genes of over 45,000 base pairs (bp), such as c-Ki-*ras* (refs 10 and 11), this potential problem is especially troubling; statistical analysis shows that even for DNA samples with an average fragment length of 100,000 bp, the c-Ki-*ras* gene would be intact in only one third of the fragments containing any part of this gene. Second, solid tumour specimens usually contain significant numbers of non-neoplastic stromal and inflammatory cells which would dilute any positive signal from the neoplastic cells in the transfection assay.

In the studies discussed here, we have investigated the prevalence and temporal occurrence of mutations of *ras* genes in colorectal cancer. To circumvent the potential problems noted above, we have used an alternative approach to detect *ras* mutations. First, biochemical methods, rather than transfection assays, were used to detect mutations at specific positions of the *ras* genes; limited amounts of DNA degradation in tumours do not affect the sensitivity of these biochemical methods. Second, we analysed only tumour areas which contained a high proportion of neoplastic cells. In many cases, this necessitated preparation of DNA from cryostat sections of tumours, in which the proportion and type of neoplastic and non-neoplastic cells could be accurately assessed.

Hybridization analysis

We initially examined haematoxylin and eosin (H & E) stained slides of colorectal cancers from 50 patients. Twenty-seven tumours were selected that showed large areas of neoplastic cells with relatively low numbers of contaminating non-neoplastic cells from either inflammation or stromal proliferation. DNA was extracted from frozen tissue samples of these tumours and analysed for the presence of *ras* mutations using oligonucleotide probes.

The sensitivity of oligomer hybridization assays can be significantly enhanced through selective amplification of short, genomic DNA fragments containing the coding sequence of interest. A technique for selective amplification of such regions was recently described^{12,13}. The amplification protocol and the oligomers used for priming amplification and detecting *ras* mutations are described in detail in ref. 14. In brief, DNA preparations from the tumours were incubated with priming oligomers complementary to sequences upstream and downstream of the codon to be tested in the presence of the Klenow fragment of DNA polymerase. After repeated cycles of denaturation, annealing with the oligomers and chain elongation, amplified DNA fragments of 60–130 bp were generated. The amplified DNA preparations were then dot-blotted onto nylon filters and hybridized with radiolabelled oligomer probes. For example, for analysis of mutations at codon 12 of the c-Ki-*ras* gene, DNA preparations were amplified using oligonucleotide primers located 31 bp upstream and 76 bp downstream of codon 12, resulting in amplified fragments of 108 bp. The amplified DNA was then divided into aliquots and dot-blotted onto seven nylon filters. Each filter was hybridized to either an individual oligomer specific for the normal glycine codon (CCTACGCCACCAGCTCCAAC, where the anti-codon 12 sequence is underlined), or to one of six oligomers specific for mutations at codon 12 (the same 20-mer but with the anti-codon sequences ACA (cysteine), ACT (serine), ACG (arginine), AAC (valine), ATC (aspartic acid) or AGC (alanine)). An example of the results of such an experiment is presented in Fig. 1. The other probes used included oligomers specific for every possible amino-acid substitution at codons 12 and 61 of the c-Ha-*ras*, c-Ki-*ras* and N-*ras* genes, as well as all possible substitutions at codon 13 of the N-*ras* and c-Ki-*ras* genes¹⁴. Although mutations induced *in vitro* at other positions of the *ras* genes can activate them in the NIH 3T3 assay¹⁵, the oligomer panel we used included probes specific for all *ras* mutations at positions known to have mutated *in vivo* in either human or animal tumours^{1–3,16}.

The patients and tumours examined in this study are described in Table 1. Nine tumours were identified with mutations at codon 12 of the c-Ki-*ras* gene. These mutations resulted in replacement of glycine (codon GGT) with aspartic acid (codon GAT, 5

§ To whom correspondence should be addressed.

Fig. 1 Dot-blot testing of colorectal carcinomas for mutations at codon 12 of the c-Ki-ras gene (K12). Amplified DNA preparations from the 27 tumour specimens (grids 1–27), selected normal colonic mucosa specimens (grids 28–33) and cell lines with known c-Ki-ras gene mutations (grids 34 and 35) were spotted onto nylon filters. Grid numbers and the corresponding amplified DNA preparations were: 1, S-7; 2, S-15; 3, S-16; 4, S-20; 5, S-22; 6, S-27; 7, S-32; 8, S-42; 9, S-43; 10, S-45; 11, S-46; 12, S-89; 13, S-92; 14, S-93; 15, S-94; 16, S-95; 17, S-96; 18, S-97; 19, S-98; 20, S-99; 21, S-100; 22, S-101; 23, S-102; 24, S-103; 25, S-104; 26, S-105; 27, S-106; 28, S-27 normal tissue; 29, S-94 normal tissue; 30, S-100 normal tissue; 31, S-101 normal tissue; 32, S-102 normal tissue; 33, S-106 normal tissue; 34, OVCA (Asp at K12 and no normal glycine allele¹²); 35, SW480 (Val at K12 and no normal glycine allele^{10,12}); 36, no sample loaded. Each filter was hybridized to a synthetic oligonucleotide probe specific for the amino acid at codon 12 of the c-Ki-ras gene indicated to the right. Thus, the 'Gly' probe (CCTACGCCACCAGCTCCAAC) detected normal c-Ki-ras genes with glycine at codon 12, where the 'Cys' probe (CCTACGCCACAAGCTCCAAC) detected mutated c-Ki-ras genes with cysteine substituted at codon 12. The variation in intensity in the glycine probe hybridization signals between different tumour samples reflected differences in the efficiencies of the *in vitro* amplification process.

Methods. Tumour specimens were obtained from patients who had undergone surgery at the Johns Hopkins Hospital and DNA was isolated from frozen specimens as previously described¹⁷. Amplified DNA preparations were prepared as in refs 12 and 13; 20 rounds of amplification were carried out. The DNA amplified *in vitro* (~0.01 µg) was spotted onto each filter. Probe preparation, pre-hybridization and hybridization were as in ref. 14. Post-hybridization filter washes were: 2×SSPE (1×SSPE is 10 mM sodium phosphate pH 7.0, 0.18 M NaCl, 1 mM EDTA), 0.1% SDS for 5 min at room temperature; 5×SSPE, 0.1% SDS for 5 min at 63 °C; hybridization buffer without Denhardt's solution or salmon sperm DNA for two brief washes at room temperature; and a final wash for 1 h in this modified hybridization buffer at 59–60 °C. The filters were exposed to Kodak XAR film at –70 °C using intensifying screens.

cases), valine (codon GTT, 2 cases), serine (codon AGT, 1 case) and cysteine (codon TGT, 1 case) (Table 1). In six of these nine cases (S-22, 27, 100, 101, 102 and 106), the hybridization signal from the oligomer recognizing the mutated codon was relatively strong (Fig. 1). In the other three tumours (S-7, 16 and 97), the hybridization signal with the mutated oligomer was relatively weak and cannot be seen in the autoradiograph shown in Fig. 1; these three cases will be discussed in detail later. A mutation at codon 61 of the c-Ki-ras gene was detected in tumour S-105; this mutation replaced the normal glutamine (codon CAA) with histidine (codon CAC). One tumour (S-98) was found to have a mutated N-ras gene. The mutation in S-98 replaced the normal glycine (codon GGT) at position 12 with cysteine (codon TGT). DNA was prepared from normal colonic mucosa of all patients exhibiting mutations in their cancers; no mutations were detected

in the normal DNA (examples shown in Fig. 1), showing that the mutations occurred somatically. In addition, no amplification of the c-Ki-ras, N-ras or c-Ha-ras genes was found in any of the tumours studied, as assessed by Southern blot analyses using cloned DNA probes for these genes (data not shown).

To confirm the results obtained with the dot-blot analysis, DNA (unamplified) from the tumours was cleaved with *Eco*RI, separated by electrophoresis through an agarose gel and hybridized with the specific oligomer giving a positive signal in the dot-blot assay. The results confirmed the dot-blot assays. In the examples shown in Fig. 2, *Eco*RI restriction fragments of 6.7 kb (containing the exon coding for amino acid 12 of c-Ki-ras) hybridized with both the normal c-Ki-ras oligomer (Fig. 2A and C) and the mutation-specific oligomers (Fig. 2B and D).

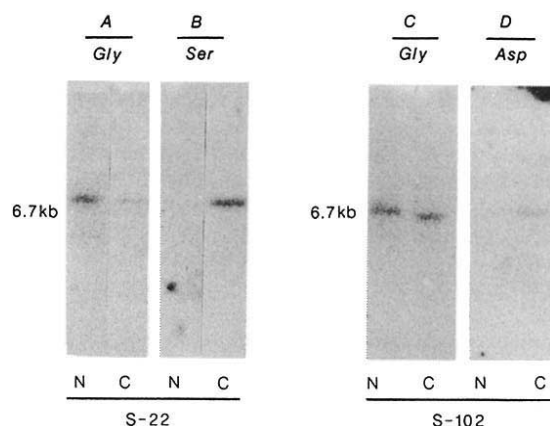


Fig. 2 Detection of mutated c-Ki-ras genes in tumours S-22 and S-102 by gel hybridization assay. Autoradiographs of dried gels containing DNA isolated from either normal colonic mucosa (N) or colorectal tumours (C) are shown. The DNA was digested with *Eco*RI and hybridized to the ³²P-labelled synthetic oligomer probe indicated: the normal c-Ki-ras codon 12 oligomer recognizing glycine (Gly) in A and C; and the mutation-specific oligomers recognizing serine (Ser) in B or aspartic acid (Asp) in D. The 6.7-kb *Eco*RI fragment contains the exon coding for amino acid 12 of the c-Ki-ras gene.

Methods. DNA was cleaved with *Eco*RI, separated by electrophoresis through agarose gels and hybridized *in situ* to ³²P-labelled synthetic oligonucleotide probes. Preparation of the oligomers, radiolabelling, hybridization, and autoradiography were as described in refs 14 and 29, and Fig. 1 legend.

Histological analysis

As noted above, tumour samples S-7, S-16 and S-97 displayed a weak hybridization signal with the mutation-specific c-Ki-ras 12 probe compared to the signal of the normal probe. There were three potential explanations for this weak hybridization signal: first, the hybridization signal with the mutation-specific probe was due to non-specific background hybridization; second, the tumour tissue was heavily infiltrated with non-neoplastic cells, thus diluting the signal from the neoplastic cells; or third, mutations were present in only a portion of the neoplastic cells. To distinguish between these possibilities, DNA was purified from cryostat sections of these three tumours, as described in refs 17 and 30. In this way, the DNA contributions from neoplastic and non-neoplastic cells in the preparations could be accurately determined by histological analysis of serial sections.

A piece of tissue from tumour S-16 was frozen in OCT compound for cryostat sectioning analysis (see Fig. 3 legend). Sections 1, 51 and 100 were stained with H & E and the remainder of the sections were used to prepare DNA. Histological analysis revealed that ~60% of the cells in the cryostat sections were neoplastic, the remainder being stromal or inflammatory (Fig. 3A). DNA analysis showed that the hybridization signal from the oligomer specific for the aspartic acid mutation was approximately one third that obtained with the glycine-specific oligomer (Fig. 4, lane c). This was consistent with the histological analysis, under the assumption that the neoplastic cells contained one mutant allele and one normal allele, whereas normal cells contained two normal alleles.

Histological analysis of tumour S-97 showed that the distribution of neoplastic cells in this specimen was non-uniform (Fig. 3B). The frozen block of tumour shown was partitioned into two areas of different neoplastic cell content; histological analysis showed that approximately 70% and 15% of cells from areas I (Fig. 3C and C') and II (Fig. 3D and D'), respectively, were neoplastic. DNA analysis showed that the hybridization signal from the aspartic-acid-specific oligomer was approximately fivefold more intense for the DNA prepared from area I (Fig. 4, lane g) compared to DNA from area II (Fig. 4, lane h), but the hybridization signal from the glycine-specific oligomer was approximately equal in intensity for the two areas (compare Fig. 4, lanes g and h). These data suggested that the great majority of the neoplastic cells in tumour S-97 contained the aspartic acid mutation and that the weak signal seen when DNA from the whole tumour specimen was hybridized to the mutated oligomer (Fig. 1 and Fig. 4, lane f) was due to contaminating non-neoplastic cells.

The histological analysis of tumour S-7 was particularly interesting. Sections through this tumour revealed that the cancer had apparently arisen from a villous adenoma, a benign colorectal neoplasm that is a precursor of colorectal cancer¹⁸. The block of tumour shown in Fig. 3E was divided into areas containing only adenoma (area I) and an adjacent area predominantly composed of malignant cells (carcinoma, area II). DNA was purified from cryostat sections from these areas (Fig. 3F and G, respectively). DNA from area II (carcinoma) hybridized strongly with the aspartic-acid-specific oligomer (Fig. 4, lane n), whereas DNA from area I (adenoma) did not (lane m). Hence, the aspartic acid mutation was present in the carcinoma, but was not present in the adenoma from which the carcinoma arose.

Adenoma-carcinoma sequence

The studies of tumour S-7 suggested that *ras* mutations might arise during the development of a carcinoma from a pre-existing adenoma. This observation prompted us to determine the temporal occurrence of *ras* mutations in other tumours containing regions of adenoma and carcinoma. Although many carcinomas arise from pre-existing adenomas, the adenomatous elements have often been replaced by carcinoma by the time cancer is

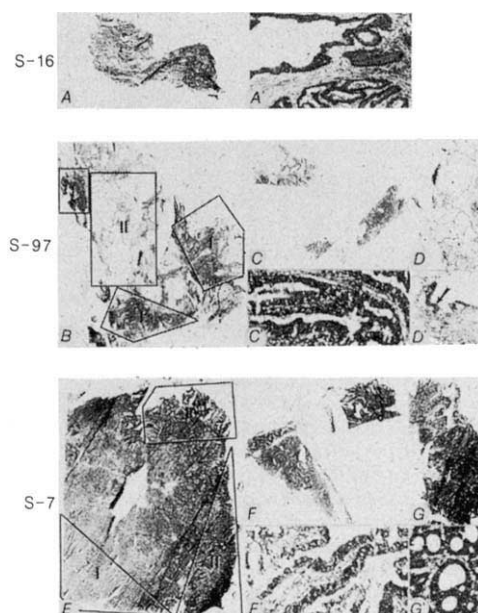


Fig. 3 Histological analysis of tumours S-16, S-97, and S-7. Tissue from tumours S-16, S-97 and S-7 was frozen in OCT compound and 100 serial 6- μ m sections were cut with a cryostat. Sections 1, 51 and 100 were stained with H & E for histological analysis, and the remainder were used to prepare DNA. A, Tumour S-16. Photographs of H & E-stained cryostat section 51 from this adenocarcinoma are shown in A (magnification $\times 2.5$) and A' ($\times 80$). B–D, Tumour S-97. A photograph of an H & E cryostat section from this mucinous adenocarcinoma is shown in B. The frozen block of tumour was cleaved using a scalpel into the regions indicated in B. The areas labelled I were selected as heavily enriched with carcinoma cells, whereas area II was selected as containing mucinous carcinoma with little malignant epithelium. Areas I and II were then refrozen in OCT compound in separate blocks and 100 serial 6- μ m sections were prepared from each block. H & E-stained cryostat section 51 from area I is shown in C and C' ($\times 2.5$ and $\times 80$, respectively); cryostat section 51 from area II is shown in D ($\times 2.5$) and D' ($\times 80$). The arrow in D' points to a region of malignant cells. E–G, Tumour S-7. A piece of tissue from tumour S-7 was frozen as described; H & E-stained section 51 from this block of tumour is shown in E. The block of tumour was then partitioned into the areas indicated in E. Area I was selected to be composed exclusively of adenoma, whereas area II was selected to be composed of carcinoma. One hundred serial 6- μ m sections were then cut from each of these two areas. H & E-stained section 51 from area I (adenoma) is shown in F and F', whereas section 51 from area II (carcinoma) is shown in G and G' ($\times 2.5$ and $\times 80$, respectively).

diagnosed¹⁸. Review of the histological sections from the 27 cancers examined in this study revealed that 8 of these tumours (S-7, S-22, S-32, S-96, S-100, S-101, S-102 and S-105) still contained residual adenoma. The cryostat sectioning technique was used to prepare DNA from the five tumours (in addition to S-7) in which *ras* mutations had been previously identified (Table 1).

Representative results of the dot-blot analysis for tumour specimens S-22, S-100 and S-101 are shown in Fig. 5. DNA prepared from normal colonic mucosa of patient S-22 (Fig. 5, lane a) hybridized to the normal glycine-specific oligomer, but not to either the serine-specific or valine-specific oligomers. In contrast, DNA isolated from the unfractionated tumour specimen (lane b), and cryostat sections of carcinoma (lane c) or adenoma (lane d) hybridized to both the glycine-specific oligomer and the serine-specific oligomer. The glycine-specific oligomer hybridized strongly to the DNA prepared from the adenoma, but only weakly to DNA prepared from the carcinoma. Because the DNA from both the adenoma and carcinoma was isolated from tumour regions that contained >90%

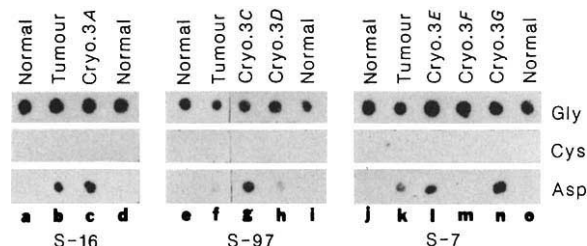


Fig. 4 Results of c-Ki-ras gene mutation analysis of DNA from cryostat sections of tumours S-16, S-97 and S-7. Autoradiographs of dot-blots of amplified DNA preparations hybridized to c-Ki-ras oligomers specific for glycine (Gly), cysteine (Cys), or aspartic acid (Asp) at codon 12 are shown. Dots a-d were prepared from DNA isolated from tissue specimens of patient S-16: a, d, normal colonic mucosa; b, the unfractionated tumour specimen; c, cryostat sections of the tumour block shown in Fig. 3A. Dots e-i were prepared from tissue specimens of patient S-97: e, i, normal colonic mucosa; f, unfractionated tumour specimen; g, h, cryostat sections of the tumour blocks displayed in Fig. 3C and D, respectively. Dots j-o contain DNA isolated from the following tissue specimens of patient S-7: j, o, normal colonic mucosa; k, the unfractionated tumour specimen; l, m, n, cryostat sections of tumour S-7 shown in Fig. 3E, F and G respectively.

Methods. DNA was prepared from normal colon or unfractionated tumour specimens by digesting ~200 mg of tissue with proteinase K and SDS, followed by extraction with phenol and chloroform¹⁷. DNA was prepared from cryostat sections of tumour blocks as follows: 100 serial 6- μ m sections were cut from the blocks with a cryostat; section 1, 51, and 100 were used for histological analysis; the remaining 97 sections were pooled and DNA purified from them as described in ref. 17. Purified DNA was amplified *in vitro* and was used for hybridization to synthetic oligonucleotide probes as in ref. 14 and Fig. 1 legend.

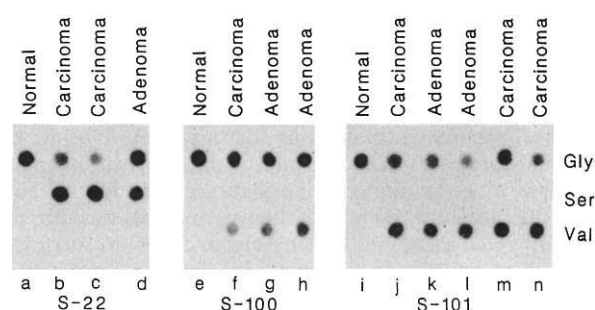


Fig. 5 Oligomer hybridization to DNA from cryostat sections of tumours containing regions of adenoma and carcinoma. Autoradiographs of dot-blots of amplified DNA preparations hybridized to oligomers specific for glycine (Gly), serine (Ser), and valine (Val) at codon 12 of the c-Ki-ras gene are shown. Dots a-d contain DNA isolated from tissue specimens of patient S-22: a, normal colonic mucosa; b, the unfractionated tumour specimen; c, d, cryostat sections containing >90% carcinoma and >90% adenoma, respectively. Dots e-h contain DNA isolated from tissue samples of patient S100: e, normal colonic mucosa; f, the unfractionated tumour specimen; g, h, cryostat sections from tumour regions each containing >90% adenoma. Dots i-n contain DNA isolated from tissue specimens of patient S-101: i, normal colonic mucosa; j, the unfractionated tumour specimen; k, l, cryostat sections of tumour regions each containing >90% adenoma; m, n, cryostat sections from tumour regions containing ~30% carcinoma and ~60% carcinoma, respectively; the remaining cells were stromal or inflammatory. DNA preparation and hybridization analysis were performed as in Fig. 4 legend.

neoplastic cells, these results suggested that normal (glycine) and mutated (serine) c-Ki-ras alleles were present in the adenoma, but that many of the carcinoma cells had lost the normal allele. However, this conclusion must be regarded as tentative, as it is difficult to quantify allele dosage by comparison of the hybridization intensities of different probes.

Tumour specimens S-100 and S-101 had previously been noted to contain a mutation that replaced glycine with valine at position 12 of the c-Ki-ras gene. In both cases, the adenoma cells from these tumours contained the same valine mutation present in the carcinoma cells (adenoma: Fig. 5, lanes g, h, k and l; carcinoma: 5, lanes f, j, m and n). Similar analysis of tumours S-102 and S-105 revealed that the adenoma from each of these two tumours contained the *ras* mutation originally identified in the tumour specimen (data not shown). Therefore, of the six tumours containing *ras* mutations and regions of adenoma and carcinoma, five were found to have the mutation in both the adenoma and the carcinoma (Table 1). Only tumour S-7 exhibited a *ras* mutation in the carcinoma which was not present in the adenoma.

Conclusions

In the studies described here we have shown the following. (1) Mutations of the *ras* gene were present in 11 of the 27 colorectal cancers examined; all mutations occurred somatically. (2) Ten of the 11 mutations were in the c-Ki-ras gene, and 9 of these were at codon 12 of this gene. In one tumour, a mutation of codon 12 of the N-ras gene was found. (3) In five of six cancers with *ras* mutations, the same mutation was found in the adenomatous (benign) and carcinomatous (malignant) regions of the tumour, suggesting that the mutation preceded the development of malignancy.

The incidence of mutations in *ras* genes in the colorectal tumours studied was significantly higher than that reported using

NIH 3T3 transfection assays in the three other types of solid tumours that have been extensively studied by other investigators: only 6% of transitional cell carcinomas of the bladder⁷ and none of the adenocarcinomas of the breast⁸ or stomach⁹ examined contained *ras* gene mutations. There are at least two potential explanations for these differences. First, it is possible that colon tumours arise through different pathogenic mechanisms than breast, stomach or bladder tumours, with *ras* mutations occurring frequently only during the development of colorectal tumours. Second, we used methods specifically designed to detect mutations that might not be detected by NIH 3T3 transfection assays of solid tumour specimens. It is possible that these methods might allow detection of a higher frequency of *ras* mutations in other tumours.

Nine of the 11 mutations detected were at position 12 of the c-Ki-ras gene. Mutations in *ras* genes in five human colon tumour cell lines (SW-480¹⁰, SK-CO1¹⁹, SW1116²⁰, SW1398²⁰ and KMS-4²¹) have also been localized to codon 12 of the c-Ki-ras gene, and one rectal tumour cell line (7060) has been reported to have a mutation at position 61 of the N-ras gene²². Thus, mutations at position 12 of the c-Ki-ras gene are either much more likely to occur in colorectal cancers than other mutations in *ras* genes, or are more likely to exert a selective growth advantage when they do occur.

One of the most intriguing observations made in this study was the detection of *ras* mutations in both the adenoma and the carcinoma of five of six tumours containing admixtures of these two types of neoplastic cells. There is abundant evidence suggesting that most colorectal cancers arise from pre-existing adenomas^{18,23-26}. Our studies of tumours containing regions of adenoma and carcinoma suggest that *ras* mutations usually precede the development of a carcinoma from a pre-existing adenoma in most cases and may therefore be involved in the early stages of human colorectal tumorigenesis. This conjecture

Table 1 Classification of patients and tumours analysed for *ras* mutations

Specimen	Age (yr)	Sex	Dukes' category*	Location of tumour	<i>ras</i> mutations†	Amino acid at mutated codon
S-7	68	F	C	Recto-sigmoid	K12	Asp
S-15	48	F	C	Sigmoid	—	—
S-16	61	F	A	Descending	K12	Asp
S-20	68	F	C	Ascending	—	—
S-22	68	F	C	Sigmoid	K12	Ser
S-27	80	F	A	Sigmoid	K12	Asp
S-32	43	F	A	Sigmoid	—	—
S-42	76	F	B	Cecum	—	—
S-43	63	F	C	Sigmoid	—	—
S-45	80	F	A	Rectum	—	—
S-46	59	F	B	Ascending	—	—
S-89	71	F	B	Ascending	—	—
S-92	75	F	B	Transverse	—	—
S-93	76	M	C	Ascending	—	—
S-94	62	M	A	Sigmoid	—	—
S-95	64	M	C	Transverse	—	—
S-96	52	M	A	Recto-sigmoid	—	—
S-97	75	M	B	Sigmoid	K12	Asp
S-98	77	M	C	Cecum	N12	Cys
S-99	58	M	C	Cecum	—	—
S-100	45	M	C	Rectum	K12	Val
S-101	73	M	B	Descending	K12	Val
S-102	67	M	C	Sigmoid	K12	Asp
S-103	67	M	C	Ascending	—	—
S-104	62	M	B	Ascending	—	—
S-105	36	M	C	Rectum	K61	His
S-106	44	M	C	Transverse	K12	Cys

* Dukes' A, tumour confined to the muscularis propria; Dukes' B, tumour extended through the muscularis propria; Dukes' C, tumour was metastatic to regional lymph nodes.

† K12, mutation at codon 12 of c-Ki-ras gene; K61, mutation at codon 61 of c-Ki-ras gene; N12, mutation at codon 12 of N-ras gene; —, no mutation identified.

is consistent with the fact that c-Ki-ras mutations were found not only in advanced colorectal cancers (Dukes' B and C tumours) but also in two cancers detected at a relatively early stage, before invasion through the muscularis propria occurred (Dukes' A; see Table 1). Studies of animal tumours provide a precedent for this finding. Balmann and colleagues found that the same c-Ha-ras gene mutation was present in both the benign skin papillomas and malignant carcinomas induced by dimethylbenzanthracene and tumour promoters²⁷. Moreover, Aaronson and co-workers have demonstrated that three of ten spontaneous hepatic adenomas of mice contained activated c-Ha-ras genes²⁸.

Note that the adenomas we studied represented a selected group; only adenomas in tumours that had progressed to the malignant state were evaluated. Thus, these studies suggest two possible hypotheses for the role of *ras* gene mutations in colonic tumour progression. Mutations in *ras* genes may occur frequently in adenomas, but only a small fraction of those adenomas with a mutation progress to malignancy. Alternatively,

ras gene mutations may occur infrequently in adenomas, but those adenomas with *ras* mutations may have a high probability of progression to the malignant state. Finally, our findings in one tumour (S-7) suggest that the genetic step promoting the adenoma-carcinoma sequence in some cancers may be a mutation in the c-Ki-ras gene. Our data do not prove this hypothesis, because the mutation in tumour S-7 could have occurred after the transition from adenoma to carcinoma. Similar studies of additional tumours, both benign and malignant, should help to define further the function and temporal occurrence of these mutations in neoplastic development of the colon.

We thank A. Preisinger, M. E. Bogaard and H. van den Elst for technical assistance; P. White for preparation of the manuscript; and Drs B. Nelkin, S. Baylin, O. M. Colvin and S. Desiderio for helpful comments on the manuscript. This work was supported by the Clayton Fund, a grant from the Netherlands Cancer Foundation KWF (MVdV), NIH training grant GM07309 (E.R.F.), and NIH grant CA35494 (B.V.).

Received 8 January; accepted 16 March 1987.

- Land, H., Parada, L. F. & Weinberg, R. A. *Science* **222**, 771-778 (1983).
- Cooper, G. M. *Science* **218**, 801-806 (1982).
- Barbacid, M. *Trends Genet.* **2**, 188-192 (1986).
- Land, H., Parada, L. F. & Weinberg, R. A. *Nature* **304**, 596-602 (1983).
- Ruley, H. E. *Nature* **304**, 602-606 (1983).
- Newbold, R. F. & Overrell, R. W. *Nature* **304**, 648-651 (1983).
- Fujita, J. et al. *Proc. natn. Acad. Sci. U.S.A.* **82**, 3849-3853 (1985).
- Kraus, M. H., Yuasa, Y. & Aaronson, S. A. *Proc. natn. Acad. Sci. U.S.A.* **81**, 5384-5388 (1984).
- Sakamoto, H. et al. *Proc. natn. Acad. Sci. U.S.A.* **83**, 3997-4001 (1986).
- Capon, D. J. et al. *Nature* **304**, 507-513 (1983).
- Shimizu, K. et al. *Nature* **304**, 497-500 (1983).
- Saiki, R. et al. *Science* **230**, 1350-1353 (1985).
- Mullis, K. B. & Faloona, F. A. *Meth. Enzym.* (in the press).
- Verlaan-de Vries, M. et al. *Gene* **50**, 313-320 (1986).
- Fasano, O. et al. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4008-4012 (1984).
- Bos, J. L. et al. *Nature* **315**, 726-730 (1985).
- Goelz, S. E., Hamilton, S. R. & Vogelstein, B. *Biochem. biophys. Res. Commun.* **130**, 118-125 (1985).

- Sugarbaker, P. H., Gunderson, L. L. & Wittes, R. E. in *Cancer: Principles and Practice in Oncology* (eds DeVita, V. T., Hellman, S. & Rosenberg, S. A.) 795-836 (Lippincott, Philadelphia, 1985).
- Winter, E., Yamamoto, F., Almoguera, C. & Perucho, M. *Proc. natn. Acad. Sci. U.S.A.* **82**, 7575-7579 (1985).
- Bos, J. L. et al. *Nucleic Acids Res.* **14**, 1209-1217 (1986).
- Yuasa, Y. et al. *Jap. J. Cancer Res.* **27**, 901-907 (1986).
- Yuasa, Y., Reddy, E. P., Rhim, J. S., Tronick, S. R. & Aaronson, S. A. *Jap. J. Cancer Res.* **77**, 639-647 (1986).
- Gilbertson, V. A. *Cancer* **34**, 936-939 (1974).
- Eide, T. J. *Int. J. Cancer* **38**, 173-176 (1986).
- Muto, T., Bussey, H. J. R. & Morson, B. C. *Cancer* **36**, 2251-2270 (1975).
- Shinya, H. & Wolff, W. I. *Ann. Surg.* **190**, 679-683 (1979).
- Quintanilla, M., Brown, K., Ramsden, M. & Balmann, A. *Nature* **322**, 78-80 (1986).
- Reynolds, S. H. et al. *Proc. natn. Acad. Sci. U.S.A.* **83**, 33-37 (1986).
- Bos, J. L. et al. *Nucleic Acids Res.* **12**, 9155-9163 (1984).
- Fearon, E. R., Feinberg, A. P., Hamilton, S. R. & Vogelstein, B. *Nature* **318**, 377-380 (1985).

Detection of high incidence of K-*ras* oncogenes during human colon tumorigenesis

Kathleen Forrester, Concepcion Almoguera, Kyuhyung Han[†], William E. Grizzle^{*} & Manuel Perucho[‡]

Department of Biochemistry, State University of New York at Stony Brook, Stony Brook, New York 11794, USA

^{*}Department of Pathology, University of Alabama at Birmingham and the Birmingham Veterans Administration Medical Center, Birmingham, Alabama 35294, USA

*RNAse A mismatch cleavage analysis of 66 primary human colon tumours reveals a high incidence of K-*ras* genes with mutations at position 12. No apparent correlation was found between the presence of mutant oncogenes and the degree of invasiveness of the tumours but evidence for *ras* mutational activation in premalignant tissue was obtained.*

INCREASING evidence indicates that the products of cellular proto-oncogenes are involved in the regulation of cell growth in both normal and pathological tissue¹. The *ras* genes and their products are among the best studied of the cellular oncogenes. The highly conserved *ras* gene family encodes closely related proteins (p21^{ras}) which are localized in the plasma membrane, bind guanine nucleotides and exhibit weak GTPase activity²⁻⁸. It is thought that *ras* proteins are involved in the process of signal transduction across the cell membrane^{9,10}.

Three members of the mammalian *ras* gene family (c-K-*ras*, c-H-*ras* and N-*ras*)¹¹⁻¹⁶ acquire oncogenic activity in cultured established cell lines by somatic mutations resulting in the substitution of single amino acids at particular positions in the p21 proteins¹⁷⁻²³.

The role of *ras* oncogenes in mammalian tumorigenesis has been well documented in various carcinogen-induced animal tumour model systems²⁴⁻²⁸. The carcinogen-specific and reproducible activation of *ras* oncogenes observed in them provides strong evidence that the activation by somatic mutation of *ras* oncogenes is involved in initiation of carcinogenesis. Similar studies are not possible in humans. Although activated *ras* oncogenes have been found in a significant percentage of human tumours (see refs 10, 29 and 30 for reviews), it has not been demonstrated that *ras* oncogene activation causes pathogenesis of human cancer. In contrast with the frequent, reproducible, and therefore predictable, mutational activation of *ras* oncogenes in some animal model systems, the incidence of mutant *ras* oncogenes in human tumours is difficult to assess with the available information¹⁰.

We have recently developed a new method for detecting and characterizing single base substitutions in transcribed genes, based on the ability of RNAse A to recognize and cleave single-base mismatches in RNA-RNA duplexes³¹. Total cellular RNA is hybridized to homogeneously labelled, antisense RNA molecules synthesized with the SP6 *in vitro* transcription system³². Digestion with RNAse A cleaves the labelled RNA probe into fragments of defined size that can be analysed by gel electrophoresis. Because the size of these fragments is dependent on the position of the mismatches in the RNA duplexes, the mutations can be localized in the gene transcripts. In addition, this method can be used to estimate the expression levels of both normal and mutant alleles in the same cell, even when they differ by only a single nucleotide³¹.

We have applied the RNAse A mismatch cleavage method to the diagnostic detection of point mutations in the first coding exon of the c-K-*ras* gene in a large panel of primary human colon tumours. Mutant genes at codon 12 were present and

expressed in about 40% of these tumours whereas mutations at other positions of the first coding exon were not detected. No apparent correlation was found between the presence of mutant oncogenes and the stage of progression of the tumours. A high frequency of *ras* oncogenes was detected in premalignant villous adenomas and in carcinomas originating in villous adenomas. The simultaneous presence of c-K-*ras* and N-*ras* oncogenes was found in a colon adenocarcinoma and in a pre-invasive villous adenoma.

K-*ras* oncogenes in colon tumours

Total cellular RNA was prepared from frozen tumours and hybridized to an antisense RNA probe complementary to the coding strand of the first coding exon of the c-K-*ras* proto-oncogene. The hybridized RNA was digested with RNAse A and analysed by denaturing polyacrylamide gel electrophoresis (PAGE). The RNA probe that we have used spans most of the c-K-*ras* first coding exon, from the *Sau*3A site to the 5' splicing acceptor site (located at nucleotide positions +102 and -11 relative to the AUG initiation codon, respectively) and about 180 additional nucleotides of upstream intron sequences³¹. Hybridization of this antisense RNA probe to mature normal c-K-*ras* mRNA followed by digestion with RNAse A generates a protected RNA duplex of 114 base pairs (bp). When this RNA probe is annealed to c-K-*ras* transcripts with single base substitutions at codon 12, the RNA duplexes contain single base mismatches (Table 1) that when cleaved by the enzyme will generate smaller RNA fragments.

Typical results are shown in Fig. 1. RNA from T24 bladder carcinoma cell line which contains a mutant c-H-*ras* oncogene^{17,18}, as well as RNA from tumours 5 and 11 (Fig. 1a) and 27 (Fig. 1b), yielded only a major band of 114 nucleotides (nt) representing the full-length protected RNA fragment. RNA from PR371 lung carcinoma cell line which contains a c-K-*ras* oncogene with a cysteine mutation at codon 12 (ref. 23 and Table 1), yielded bands of 46 and 67 nt diagnostic of a mismatch at this position in the c-K-*ras* transcripts. These bands were also present when RNA from tumours 3, 7 and 22 (Fig. 1a), and 22, 10, 25, 28 and 23 (Fig. 1b), were hybridized to the RNA probe and digested with RNAse A, indicating the presence of mutant c-K-*ras* genes at codon 12 in these tumours.

Fragments of the c-K-*ras* genes from tumours 3 and 28 were cloned and their first coding exons were sequenced. These experiments confirmed the presence of point mutations in the c-K-*ras* gene from these tumours, which were characterized as cysteine (TGT) and serine (AGT) substitutions at codon 12 (Table 1), respectively (unpublished). These results demonstrate the applicability of the RNAse A mismatch cleavage method for the diagnostic detection in primary tumours of *ras* genes containing single point mutations.

The sensitivity of the RNAse A method for detecting single

[†] Present address: Department of Biological Sciences, Columbia University, New York, New York 10027, USA.

[‡] To whom correspondence should be sent.

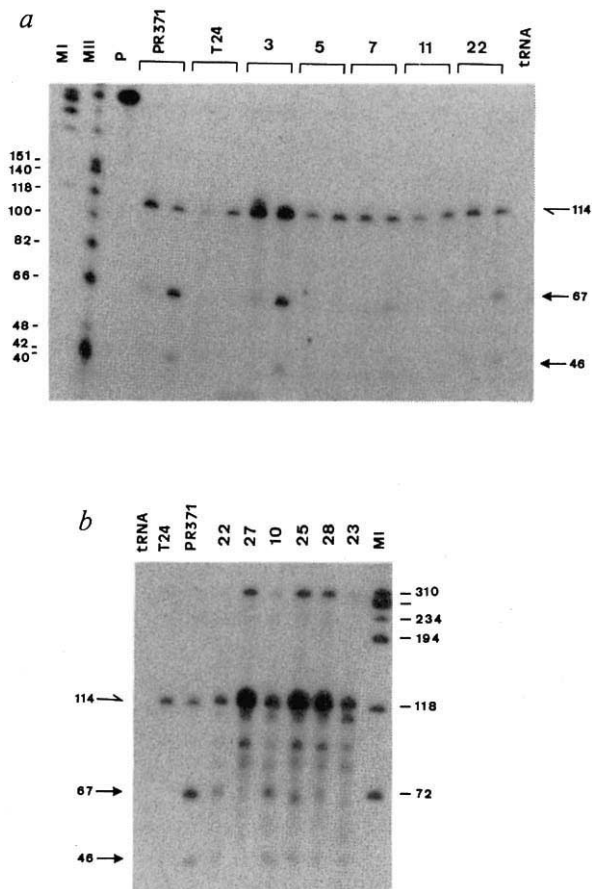


Fig. 1 RNase A mismatch cleavage analysis of c-K-ras transcripts from human colon tumours. Total cellular RNA from the tumours and tumour cell lines indicated at the top were hybridized to radioactively labelled antisense RNA corresponding to the first coding exon of the c-K-ras proto-oncogene (10^5 c.p.m.) for 16 h at 50 °C. Hybridized RNAs were digested at 30 °C with RNase A ($40 \mu\text{g ml}^{-1}$) for 10 and 60 min (from left to right) (a) or at 35 °C with $75 \mu\text{g ml}^{-1}$ of enzyme for 90 min (b) and electrophoresed through denaturing polyacrylamide gels. Amounts of RNA used ranged from 20 to 80 μg . Lane marked tRNA, only carrier transfer RNA was used during hybridization. Numbers at left (a) or right (b), indicate the position and size (in nt) of *Hae*III (MI) or *Hinf*I (MII) restriction fragments of $\phi \times 174$ DNA used as size markers. Complete arrows indicate the position and size (in nt) of the RNA fragments generated by RNase A cleavage at the mismatch of codon 12. Half arrows indicate the position and size of the protected RNA fragments corresponding to the first coding exon. P: undigested RNA probe.

Methods. RNA was prepared from cultured cells or from frozen tumours (0.2–1 g) by the guanidine HCl method as described³¹. Quality of RNAs was tested by formaldehyde-agarose gels and staining with ethidium bromide. Only samples of tumour RNA showing the presence of ribosomal 28S and 18S RNA stained bands at approximate ratios of 2:1 were used. Radioactive RNA probes were synthesized using 1 μg of linearized pAK1Ngly (ref. 31) with the SP6 polymerase *in vitro* transcription system³² as described³¹. Hybridizations and RNase A digestions were as described³¹ using gel purified RNA probes³³. RNA samples were analysed by polyacrylamide gel electrophoresis and autoradiography as described³¹.

base substitutions depends on the efficiency of cleavage at the mismatch by the enzyme^{31,33}. Thus, whereas the mismatches present in RNA hybrids of some tumour RNAs were readily cleaved by RNase A (Fig. 1a), an increase in the digestion time and temperature, and enzyme concentration, was required to obtain significant cleavage of other samples (Fig. 1b). These conditions also increased the background bands due to non-

Table 1 RNA-RNA mismatches generated by mutations at codon 12 of the human c-K-ras gene

	Glycine	Cysteine	Serine	Arginine	Valine	Aspartic	Alanine
	GGU	UGU	AGU	CGU	GUU	GAU	GCU
Glycine	5' 3'	5' 3'	5' 3'	5' 3'	5' 3'	5' 3'	5' 3'
Cysteine	5' 3'	5' 3'	5' 3'	5' 3'	5' 3'	5' 3'	5' 3'
Serine	5' 3'	5' 3'	5' 3'	5' 3'	5' 3'	5' 3'	5' 3'
Valine	5' 3'	5' 3'	5' 3'	5' 3'	5' 3'	5' 3'	5' 3'

Amino acids encoded by the normal (Gly) and mutant codon 12 of the human c-K-ras gene and the corresponding nucleotide sequence of the gene transcripts (reading 5' to 3') are indicated at the top. Nucleotide sequences of the 12th codon in antisense RNA (reading 3' to 5') corresponding to the normal or mutant c-K-ras transcripts are shown paired to the codon 12 mRNA sequences. Perfectly paired RNA-RNA duplexes are indicated by circles. All single or double mismatches are boxed. Hatched boxes represent mismatches recognized and cleaved by RNase A (ref. 31 and C.A., K.F. and M.P., unpublished data). Empty boxes, mismatches that have not been analysed. Double box, the G-U mismatch that is not recognized by the enzyme.

specific cleavage by RNase A of perfectly matched RNA duplexes, especially in A+U-rich regions.

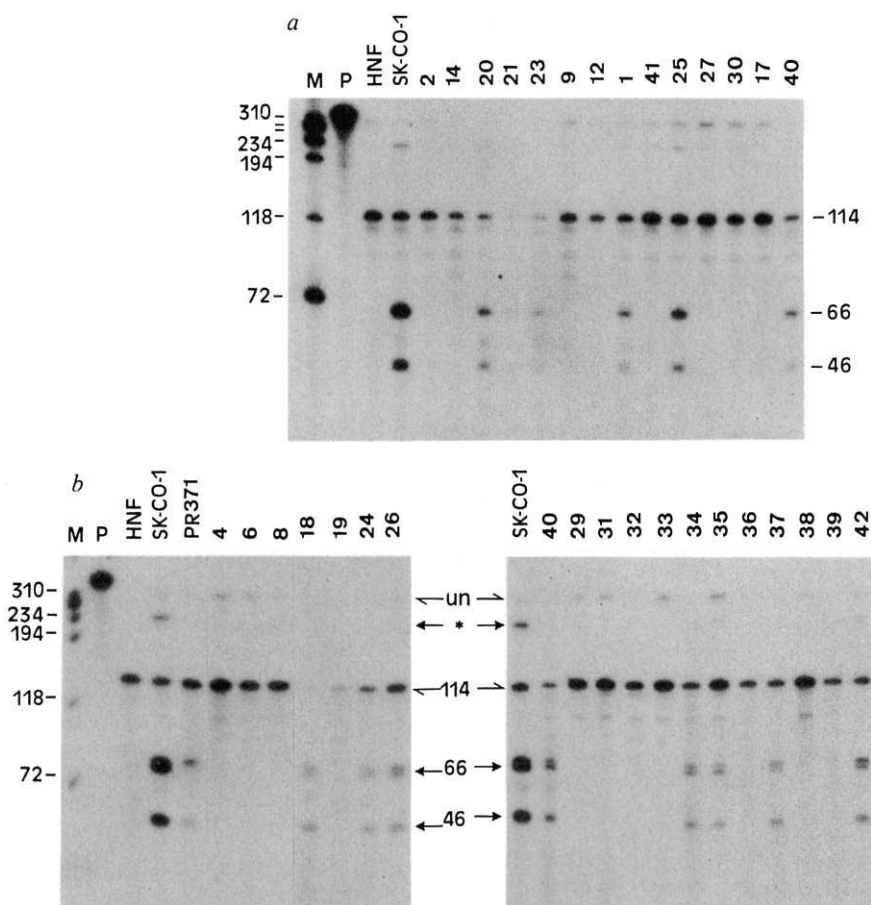
We took advantage of the finding that the G-U mismatch generated by hybridization of normal c-K-ras transcripts with the antisense RNA probe corresponding to the gene with a serine mutation at codon 12 (Table 1) is not cleaved by RNase A under the standard conditions of digestion. This is not surprising because these bases can pair in RNA-RNA duplexes³⁴. However, c-K-ras transcripts with mutations at the second position of codon 12 generate RNA hybrids containing a double mismatch when hybridized with the serine RNA probe: the G-U mismatch at the first position and another mismatch (A-C, U-C or C-C) at the second position of the triplet (Table 1). These double mismatches are now recognized and cleaved by the enzyme very efficiently.

Mutation frequency of K-ras

We screened colon tumours for the presence of mutant c-K-ras genes by hybridizing their RNAs to the serine RNA probe and digesting the resulting hybrids with RNase A (Fig. 2). RNA from normal human fibroblasts (HNF) and tumours 2, 14, 9, 12, 41, 30, 27 and 17 (Fig. 2a) and 4, 6, 8, 19, 29, 31, 32, 33, 36, 38 and 39 (Fig. 2b) yielded only a major protected RNA fragment of 114 nucleotides (together with other minor nonspecific bands). RNA from tumours 20, 21, 23, 1, 25 and 40 (Fig. 2a) and 18, 24, 26, 40, 34, 35, 37 and 42 (Fig. 2b), yielded the 66- and 46-nt bands diagnostic of the presence of c-K-ras transcripts with mutations at the second base of codon 12. RNA from SK-CO-1 colon carcinoma cell line, which contains a c-K-ras gene with a valine mutation (Table 1) at codon 12 (C.A., K.F. and M.P., manuscript in preparation), and from PR371 (Fig. 2b) was also cleaved, yielding the codon 12 mismatch-diagnostic bands. A faint band of about 300 nt was present in most cases (indicated as 'un' in Fig. 2b), representing the unspliced c-K-ras transcripts (absence of this band in some samples was probably due to degradation of the RNAs). This band was also cleaved in the positive samples, yielding a sub-band of about 230 nt (indicated as '*' in Fig. 2b).

Fig. 2 Analysis of c-K-ras transcripts mutant at the second base of codon 12. Total cellular RNA from tumours and cell lines indicated at the top were hybridized to 32 P-labelled anti-sense RNA (10^5 c.p.m.) directed by pAK1Mser, for 12 h at 50 °C. Hybridized RNAs were digested at 30 °C (a) or 25 °C (b) for 60 min and analysed in denaturing polyacrylamide gels. Amounts of RNA ranged from 10 to 45 μ g. Symbols in b: un, unspliced transcript; *, cleavage product of the unspliced transcript. Other symbols are as in Fig. 1.

Methods. A plasmid directing the synthesis of antisense RNA containing a serine mutation at codon 12 of the c-K-ras gene (pAK1Mser) was constructed by subcloning into the pGEM-4 vector the Sau3A-PstI 280-base pair fragment of the c-K-ras gene isolated from colon tumour 28.



The results obtained with a series of 66 primary human colon tumours are summarized in Table 2. Twenty-six samples were positive in the RNase A mismatch cleavage assay. Of those, 8 were characterized as point mutations at the first base and 18 at the second base of codon 12, by comparing their cleavage patterns with the normal and mutant RNA probes.

In contrast with the high frequency of mutant c-K-ras genes (39%) observed with the RNase A mismatch cleavage method, only 20% (6 of 30) of the tumours that we have analysed were scored as positive in the NIH 3T3 focus formation assay (Table 2).

K-ras expression

Our studies also provided information on the expression levels of both normal and mutant c-K-ras alleles in human colon tumours. No clear differences were observed in the expression levels of the c-K-ras gene in different tumours, regardless of the presence or absence of mutations at codon 12. Comparison of the c-K-ras gene expression levels in neoplastic and normal tissue from the same individual also failed to reveal significant differences (data not shown).

Analysis of the relative expression of mutant and normal alleles in primary tumours was complicated by the presence of varying amounts of contaminating normal tissue in the tumour samples and by incomplete cleavage of some of the mismatches. However, the efficient cleavage of the double mismatch generated by hybridization of the c-K-ras transcripts with the serine RNA probe allowed an approximate estimate of the relative expression of c-K-ras genes containing mutations at the second position of codon 12 (Fig. 2). In most cases, the expression levels of mutant and normal alleles were similar (compare the relative intensities of protected and cleaved radioactive bands). SK-CO-1 RNA used as positive control was apparently cleaved more efficiently because in these cells the mutant transcripts are

two- to threefold more abundant than the normal transcripts (C.A., K.F. and M.P., in preparation).

K-ras and tumour progression

Analysis of the data of Table 2 reveals several features. No apparent correlation was found between the presence or absence of mutant c-K-ras genes and the age, sex or race of the cancer patients, or the anatomical localization of the tumours, supporting the somatic nature of *ras* gene mutational activation. In this context, in all ten cases that we have analysed, no mutations in the first coding exon of the c-K-ras gene were found in normal colon tissue adjacent to tumours positive for mutant oncogenes by the RNase A mismatch cleavage assay (data not shown).

No significant differences were found between the mutation frequency at position 12 of the c-K-ras gene and the degree of differentiation or stage of progression of the tumours. Mutant genes were found in tumours ranging from early stages of progression to very advanced stages of disease (Table 2). However, a striking difference was apparent between the position of the mutations at codon 12 of the c-K-ras gene and the degree of invasiveness of the tumours. Whereas 4 of the 8 mutations at the first position of codon 12 (NGT) were present in the gene of tumours at early stage of progression (grades O and A), 16 of the 18 mutations at the second position of the triplet (GNT) were present in the gene of more invasive tumours (grades B, C and D) (Table 2). Further studies will be necessary to determine the significance of these differences and their possible prognostic value in human colon cancer.

ras oncogenes in premalignant tumours

In terms of the histological features of the tumours, a strikingly high incidence of mutant c-K-ras oncogenes was found in tumours originating in villous adenomas and villoglandular polyps. These are relatively common types of benign tumours

Table 2 Frequency of mutations at codon 12 of the c-K-ras gene in primary human colon tumours

Tumour no.	Mutation*	Transfection assay‡	Donor§	Location	Differentiation¶	Invasiveness #	Metastases**	Comments††
1	+GNT	—	61 FW	C	M	B	0/25	
2	—	—	77 MW	S	M	B	0/16	
3	+TGT†	+	68 FW	R	M	A	0/20	(V)
4	—	ND	60 MW	S	M	A	0/18	(AP) (MF)
5	—	ND	58 MW	S	M	C	6/42	
6	—	—	77 MB	A	W/M	A	0/23	(MC)
7	+NGT	ND	64 FB	S	M	B	1/1	
8	—	—	74 MW	T	M	B	0/37	(AP)
9	—	—	76 MW	A	M	C	2/31	
10	+XGT	—	73 FW	D	M	C	1/22	
11	—	ND	64 FW	D	M	C	1/18	
12	—	ND	66 FW	A		B	0/26	(AP)
13	—	—	62 MW					
14	—	—	52 MW	S	M	B	0/1	
15	—	ND	51 FB					
16	—	ND	79 F					
17	—	ND	50 MW					
18	+GNT	—	62 FB	A	W	C	+	
19	—	—	62 MW	S	M	C	1/15	
20	+GNT	ND	66 FW	C		B	—	
21	+GNT	—	52 MB	D	W	C	1/1	
22	+XGT	—	59 FB	S	M	C	2/2	
23	+GNT	—	53 MW	A	M	C	5/30	
24	+GNT	—	51 FB	D	W	C	1/1	(MC)
25	+GNT	—	65 MW	A	M	C	11/29	(AP)
26	+GNT	+	63 MW	C	W	A	0/27	(AP)
27	—	—	63 MB	D	W/M	C	+	
28	+AGT†	+++	68 MB	S	Cl	O	—	(V)
29	—	ND	78 FB	T	M	A	0/23	(MC)
30	—	ND	77 FW	S	W	A	0/34	
31	+XGT	—	70 FB	C	M	C	4/15	(V) (MC)
32	—	—	60 FW	C	P	C	5/13	(SR)
33	—	—	61 MW	R	M/P	B	0/6	
34	+GNT	—	67 MW	R	M	C	4/12	
35	+GNT	+	67 MW	T	W	D	3/22	(MC)
36	—	ND	80 FB	A	M	B	0/19	
37	+GNT	++§	64 FW	S	M	B	0/35	
38	—	ND	72 FB	C	W/M	C	3/15	
39	—	ND	74 MW	R	W/M	C	4/13	
40	+GNT	ND	69 FW	S	M	D	+	
41	—	ND	67 MW	S	M	D	+	
42	+GNT	—	77 MW	A	W	B	0/70	(V) (MC) (AP)
43	+XGT	—	56 MW	C	W	A	0/17	(V) (MC)
44	—	ND	77 FW	S	M	A	0/29	(MC)
45	—	—	76 MW	R	M	B	0/24	
46	—	ND	65 MB	A	M	B	0/25	
47	+GNT	—	59 MW	R	M	B	0/7	
48	—	ND	81 MW	T	M	A	0/23	(AP)
49	—	ND	60 MW	T	M	B	0/23	(MC)
50	—	+++	74 MW	A	M	B	0/22	(V) (MC)
51	—	ND	67 FW	A	M	A	0/63	(MC)
52	—	ND	60 MW	S	M	C	2/46	
53	—	ND	73 MW	R	M	C	3/31	
54	—	ND	34 FW	T	P	D	4/9	
55	+GNT	ND	68 MB	S	M	B	0/24	(MC)
56	—	ND	76 MW	S	M	C	1/7	
57	—	ND	71 MW	C	M	B	0/4	
58	+GNT	ND	69 MW	A		A	0/4	(V) (MC)
59	+GNT	ND	56 MW	R	M	B	0/25	
60	+NGT	ND	75 FW	C	W	A	0/31	(V)
61	—	ND	55 FB	A	M	B	0/37	
62	—	ND	62 MW	C	M	B	0/12	
63	—	ND	65 MW	R	M	B	0/1	
64	—	ND	66 MW	C	M	A	0/13	
65	+GNT	ND	64 MW	S	M/W	C	1/5	
66	—	ND	76 FW	R	M	B	0/5	

Summary

Type	RNase A	Transfection assay	O	A	Invasiveness B	C	D	Metastases -	+
?	0	1							
NGT	8	2	1	3	1	3	0	5	3
GNT	18	3	0	2	7	7	2	9	9
Total (%)	26/66 (39)	6/30 (20)	1/1	5/13 (38)	8/23 (34)	10/21 (47)	2/4 (50)	14/37 (37)	12/25 (48)

* Presence (+) or absence (—) of single nucleotide changes at codon 12 of the c-K-ras gene detected by the RNase A cleavage method. The position of the mutated nucleotide in the normal triplet (GGT) is underlined. N: A, T or C. X: A or T. Absence of G-C transversions at the first nucleotide of codon 12 was determined by Southern blot hybridization of tumour DNAs digested with SstI⁴³. †, Mutations characterized by cloning and sequencing. ‡ Presence (+) or absence (—) of c-K-ras oncogenes as determined by the NIH 3T3 transfection assay⁴⁴. Results were considered negative only after at least two independent experiments with a total amount of DNA of at least 150 µg. Only DNA preparations giving efficiencies of more than 0.2 colonies per µg for the transfer of the thymidine kinase gene to Ltk⁺ cells in parallel transfection experiments⁴⁴ were considered. § Age, sex and race of the cancer patients. || Anatomical localization of the tumours: C, caecum; A, ascending; T, transverse; D, descending; S, sigmoid colon; R, rectum. ¶, Differentiation grade (well (W), moderate (M) or poorly (P) differentiated). Cl, carcinoma *in situ*. # Grade of invasiveness of the tumours following Duke's classification⁴⁵. O, non-invasive tumour. ** Presence (+) or absence (—) of detectable metastases at time of surgery. Number of metastatic lymph nodes/total number analysed. †† All cases were primary colon tumours from previously untreated patients. V, carcinoma originating in a villous adenoma or villoglandular polyp; AP, presence of multiple adenomatous polyps; MC, mucinous carcinoma; SR, signet ring cell carcinoma. ‡‡ Tumour positive for both c-K-ras and N-ras oncogenes (see text). §§ Transforming activity of tumour DNA in NIH 3T3 cells. Oncogene not characterized.

Frozen colon tumour specimens (1 and 9–25) were obtained from the National Cancer Institute Repository, Division of Cancer Etiology, Biological Carcinogenesis Branch, NIH, Bethesda MD 20205. Tumours 40 and 41 were obtained from the University Hospital of SUNY at Stony Brook. Frozen colon tumours and normal tissue from the same patients were obtained from the Tissue Procurement Service of the Comprehensive Cancer Center, University of Alabama at Birmingham, Alabama. Information on tumour diagnosis and patient data were obtained from the surgical histopathology reports.

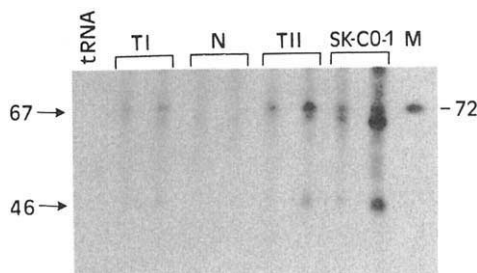


Fig. 3 RNase A mismatch cleavage analysis of c-K-ras transcripts from villous adenoma 28. RNA (30 µg) prepared from two different portions of the same tumour (T1 and TII) as well as 60 µg of RNA prepared from normal colon tissue from the same individual (N) were hybridized to ³²P-labelled antisense RNA corresponding to the normal c-K-ras first coding exon. The hybrids were digested with RNase A at 30 °C for 30 and 90 min (shown from left to right) and analysed in a denaturing polyacrylamide gel as before. RNA from the SK-CO-1 cell line (30 µg) was also analysed in parallel as positive control. Symbols are as in Fig. 1.

which emerge directly from the colonic mucosa as large superficial neoplasms and which develop malignant properties more frequently than do adenomatous polyps. Thus, these types of benign tumours often contain focal areas of conversion to carcinoma. However, such tumours are only considered malignant when infiltrating activity can be demonstrated by microscopic analysis³⁵.

Of the eight colon tumours originating in villous adenomas that we have analysed, seven contained c-K-ras genes mutant at the first coding exon, five at the first base of codon 12 and two at the second (Table 2). Interestingly, the only tumour which was scored as negative in the RNase mismatch cleavage method (tumour 50) was positive in the NIH 3T3 focus assay (Table 2), suggesting the presence of another *ras* oncogene or of a c-K-ras gene with a mutation at another position.

These eight tumours were collected at different stages of progression. Tumour 31 was metastatic at time of surgery (grade C), tumours 42 and 50 had already perforated the colon wall (grade B) and tumours 43 and 58 had infiltrated the muscularis propria (grade A). Tumours 60, 3 and 28 were obtained at very early stages of development. Histological analysis of tumour 60 revealed a villoglandular polyp containing multiple small foci of well-differentiated adenocarcinoma which were superficial with no evidence of stalk invasion. Tumour 3 was classified as a moderately differentiated adenocarcinoma, arising within a villous adenoma but the carcinoma did not invade beyond the submucosal layer. Histopathological analysis of tumour 28 revealed a villous adenoma with focal areas of carcinoma *in situ* (no invasion) and was accordingly classified as a premalignant villous adenoma.

Fragments of frozen tissue from these last two tumours (3 and 28) were resubmitted for microscopic analysis. A fragment of tumour 3 adjacent to the tissue used for both DNA and RNA preparations contained only villous adenoma tissue. Two additional distinct fragments of tumour 28 (fragments I and II) were similarly analysed. Whereas fragment I contained ~10–20% of carcinoma *in situ*, fragment II contained only villous adenoma tissue. Both fragments contained normal colon tissue in proportions varying from 10 to 30%.

Coexistence of K-ras and N-ras oncogenes

Tumour 28 was initially found to contain a mutated c-K-ras gene by the RNase A cleavage method (Fig. 1b). Parallel analysis of the transforming activity of DNA from this tumour in the NIH 3T3 focus formation assay revealed the presence of a transforming gene which was identified by Southern blotting of NIH 3T3 primary transformants as the N-ras oncogene (data

not shown). However, subsequent cloning and sequence analysis confirmed a serine mutation at codon 12 of the c-K-ras gene present in the same tumour DNA preparation that yielded N-ras NIH 3T3 transformants.

Repeated analyses by the RNase A mismatch cleavage method (Fig. 3) and the NIH 3T3 focus assay (data not shown) with different fragments of the same frozen tumour immediately adjacent to the fragments I and II previously submitted for microscopic characterization confirmed the simultaneous presence of activated N-ras and c-K-ras oncogenes in preinvasive villous adenoma 28. Similar results were obtained with tumour 37, which contained a mutant c-K-ras oncogene as detected by the RNase mismatch cleavage method (Fig. 2) and an activated N-ras oncogene as detected by the focus formation assay (Table 2 and data not shown).

Discussion

We have used a new method for the diagnostic detection of single point mutations in transcribed genes to study the frequency of c-K-ras oncogenes in a series of primary human colon tumours. Our results show that the association of somatic mutations at codon 12 of the c-K-ras gene with this type of human cancer is significantly higher than previously estimated using the NIH 3T3 transfection assay¹⁰. The explanation for this difference most probably lies in the sensitivity of these methods. The NIH 3T3 focus formation assay is particularly insensitive in the detection of activated c-K-ras genes due to the large size of this locus^{21,23}. In agreement with these results, DNA of some tumours containing mutant c-K-ras genes failed to induce foci in NIH 3T3 cells despite their ability to transfer in parallel experiments the human thymidine kinase gene (Table 2). Owing to the limitations in the sensitivity of our method and considering that we have not analysed other regions of the c-K-ras gene³⁶, or the other two characterized human *ras* oncogenes, our estimate that 40% of human colon carcinomas contain activated *ras* oncogenes may be conservative.

The RNase A mismatch cleavage method, unlike the oligonucleotide hybridization procedure^{27,37–40}, is not limited to a defined localization of the point mutations in the gene. However, we have not found mutations in the first coding exon of the c-K-ras gene at positions other than codon 12. This result argues against the hypothesis that mutations in *ras* genes are a secondary consequence of the intrinsically high mutation rate of cancer cells with no direct association with the tumorigenesis process⁴¹.

Taking into account the presence of contaminating normal tissue in the tumour samples, the information obtained on the expression levels of c-K-ras genes mutant at the second position of codon 12 suggests that the mutant allele was present and expressed in most of the tumour cells in most of the positive samples. These results support the concept that mutational activation of *ras* genes is directly involved in tumour progression by conferring a selective growth advantage on the cells in which it occurs^{10,29,30}. At the same time, the apparent presence and expression of the normal allele in these tumours suggest that loss of the normal allele is not a prerequisite for tumour development.

Although our findings on the mutational activation of the c-K-ras gene in tumours 3 and 60 and of the c-K-ras and N-ras genes in villous adenoma 28 do not exclude the possibility that the mutations in these genes may have occurred after the conversion from adenoma to carcinoma or to focal carcinoma *in situ*, we consider this possibility unlikely. Repeated microscopic examinations indicated that if carcinoma cells were present in the tumour samples that we used for preparation of DNA and RNA, their relative proportion was no greater than 10–20%. Unless the mutant *ras* alleles has undergone considerable amplification in these cells, it is unlikely that they could have been so readily isolated from genomic libraries and detected by the RNase A mismatch method and the NIH 3T3 focus assay.

Therefore, we interpret these results as indicative that *ras* mutational activation has already occurred in the benign tumour cells.

It is tempting to suggest a parallelism in *ras* oncogene activation between our results with human colon villous adenomas/carcinomas and the initiation/promotion experimental system of mouse skin papillomas/carcinomas^{24,28,42}. In this system, topical application of the carcinogen DMBA (7,12-dimethylbenzanthracene) induces the mutational activation of the c-H-*ras* gene which may give rise to the benign skin papillomas after treatment with tumour promoters. Some of these papillomas undergo further alterations and progress to malignant carcinomas.

The high frequency of mutant c-K-*ras* genes that we have found in villous adenomas and in colon carcinomas originating in villous adenomas suggests that *ras* somatic mutational activation is also involved in the development of these benign human tumours. However, the simultaneous presence of both c-K-*ras* and N-*ras* oncogenes in villous adenoma 28 and in adenocar-

cinoma 37 suggests that mutational activation of at least one of these genes was not the initial event in the development of these tumours. Therefore, mutational activation of human *ras* oncogenes might both occur and also be maintained by selection after initiation of tumorigenesis. Our experiments do not prove the simultaneous presence of both *ras* oncogenes in the same tumour cells and we cannot rule out the possible polyclonal origin of these tumours.

In conclusion, our studies support the concept that activated *ras* oncogenes are contributing in a dominant manner to the process of tumour progression and provide evidence for the frequent association of *ras* mutational activation with the early stages of human tumorigenesis.

We thank C. Lama for technical assistance and K. Sexton for help with the collection of tumour specimens. Some of the tumours used in this study were supplied by M. Viola. This work was supported by an NIH grant awarded by the National Cancer Institute and by a grant from Toyo Jozo Ltd.

Received 29 January; accepted 25 March 1987.

- Bishop, J. M. *et al. Rev. Biochem.* **52**, 301-354 (1983).
- Willingham, M. C., Pastan, I., Shih, T. Y. & Scolnick, E. M. *Cell* **19**, 1005-1014 (1980).
- Willumsen, B. M. *et al. Nature* **310**, 583-586 (1984).
- Scolnick, E. M., Papageorge, A. G. & Shih, T. Y. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5355-5359 (1979).
- Shih, T. Y. *et al. Nature* **287**, 686-691 (1980).
- McGrath, J. P., Capon, D. J., Goeddel, D. V. & Levinson, A. D. *Nature* **310**, 644-649 (1984).
- Sweet, R. W. *et al. Nature* **311**, 273-275 (1984).
- Gibbs, J. B., Sigal, I. S., Poe, M. & Scolnick, E. M. *Proc. natn. Acad. Sci. U.S.A.* **81**, 5704-5708 (1984).
- Gilman, A. G. *Cell* **36**, 577-579 (1984).
- Barbacid, M. *et al. Rev. Biochem.* (in the press).
- Ellis, R. W. *et al. Nature* **292**, 506-510 (1981).
- Der, C. J., Kravitz, T. G. & Cooper, G. M. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3637-3640 (1982).
- Parada, L. F., Tabin, C. J., Shih, C. & Weinberg, R. A. *Nature* **297**, 474-478 (1982).
- Santos, E. *et al. Nature* **298**, 343-347 (1982).
- Shimizu, K. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 2112-2116 (1983).
- Hall, A., Marshall, C. J., Spurr, N. K. & Weiss, R. A. *Nature* **303**, 396-400 (1983).
- Tabin, C. J. *et al. Nature* **300**, 143-149 (1982).
- Reddy, E. P., Reynolds, R. K., Santos, E. & Barbacid, M. *Nature* **300**, 149-152 (1982).
- Yuasa, Y. *et al. Nature* **303**, 775-779 (1983).
- Taparowsky, E., Shimizu, K., Goldfarb, M. & Wigler, M. *Cell* **34**, 581-586 (1983).
- Shimizu, K. *et al. Nature* **304**, 497-500 (1983).
- Capon, D. J. *et al. Nature* **304**, 507-513 (1983).
- Nakano, H. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 71-75 (1984).

- Balmain, A. & Pragnell, I. B. *Nature* **303**, 72-74 (1983).
- Sukumar, S., Notario, V., Martin-Zanca, D. & Barbacid, M. *Nature* **306**, 658-661 (1983).
- Guerrero, I., Calzada, P., Mayer, A. & Pellicer, A. *Proc. natn. Acad. Sci. U.S.A.* **81**, 202-205 (1984).
- Zarbl, H. *et al. Nature* **315**, 382-385 (1985).
- Quintanilla, M., Brown, K., Ramsden, M. & Balmain, A. *Nature* **322**, 78-80 (1986).
- Varmus, H. E. *et al. Rev. Genet.* **18**, 553-612 (1984).
- Marshall, C. in *RNA Tumor Viruses* 2nd Edn (eds Weiss, R., Teich, N., Varmus, H. & Coffin, J.) 487-558 (Cold Spring Harbor Laboratory, New York, 1985).
- Winter, E., Yamamoto, F., Almoguera, C. & Perucho, M. *Proc. natn. Acad. Sci. U.S.A.* **82**, 7575-7579 (1985).
- Melton, D. A. *et al. Nucleic Acids Res.* **12**, 7035-7056 (1984).
- Myers, R. M., Larin, Z. & Maniatis, T. *Science* **230**, 1242-1246 (1985).
- Tinoco, I. Jr *et al. Nature new Biol.* **246**, 40-41 (1973).
- Ashley, D. J. B. in *Histological Appearances of Tumors* 580-581 (Churchill Livingstone, London, 1978).
- Yamamoto, F. & Perucho, M. *Nucleic Acids Res.* **12**, 8873-8885 (1984).
- Bos, J. L. *et al. Nature* **315**, 726-730 (1985).
- Bos, J. L. *et al. Nucleic Acids Res.* **12**, 9155-9163 (1984).
- Guerrero, I., Villasante, A., Corces, V. & Pellicer, A. *Science* **225**, 1159-1162 (1984).
- Valenzuela, D. M. & Groffen, J. *Nucleic Acids Res.* **14**, 843-852 (1986).
- Duesberg, P. H. *Science* **228**, 669-677 (1985).
- Balmain, A., Ramsden, M., Bowden, G. T. & Smith, J. *Nature* **307**, 658-660 (1984).
- Santos, E. *et al. Science* **233**, 661-664 (1984).
- Perucho, M. *et al. Cell* **27**, 467-476 (1981).
- Sugarbaker, P. H., Gunderson, L. L. & Wittes, R. E. in *Cancer. Principles and Practice of Oncology* (eds De Vita, V. T., Hellman, S. & Rosenberg, S. A.) 814-816 (Lippincott, Philadelphia, 1985).

LETTERS TO NATURE

Possible relation between the X-ray QPO phenomenon and general relativity

Bohdan Paczyński

Princeton University Observatory, Princeton University,
Princeton, New Jersey 08544, USA

Many bright X-ray sources were recently found to undergo quasi-periodic oscillations¹⁻⁴. None of the sources is an X-ray pulsar which indicates that the accreting neutron stars have no magnetospheres. Most neutron star models have radii smaller than 3 Schwarzschild radii⁵. Between 69% and 86% of the accretion energy is liberated when matter falls from the marginally stable orbit with a radius $r_{ms} = 3r_{Sch}$ onto the neutron star surface⁶. The accretion flow through r_{ms} is known to be complicated, and a steady-state flow may not be possible when the viscosity parameter $\alpha > \alpha_{cr} \approx 0.03^{7,8}$. Here I suggest that the unsteady flow would make the boundary-layer luminosity variable, possibly giving rise to the X-ray quasi-periodic oscillation (QPO) phenomenon.

Thin disk accretion onto a black hole was a matter for many theoretical studies, with a classic paper written by Novikov and Thorne⁹. The disk accretion is subsonic, the streamlines are almost identical with the keplerian orbits beyond the marginally stable orbit, $r > r_{ms}$, and matter falls freely into the black hole at $r < r_{ms}$. Some insight into the character of accretion flow near r_{ms} was provided by the theory of thick accretion disks^{10,11}. Near

the marginally stable orbit the radial velocity of accreting matter changes from subsonic for $r > r_{ms}$ to supersonic for $r < r_{ms}$. The sonic radius is slightly smaller than r_{ms} in thick disks, as well as in thin disks with $\alpha \ll 1^{12-14}$.

Muchotrzeb⁷ noticed that the flow in thin accretion disks is complicated for $\alpha > \alpha_{cr} \approx 0.03$. In fact she could not find stationary solutions in this case. Subsequent studies revealed that while the critical (sonic) point of the flow was of a saddle type for $\alpha < \alpha_{cr}$ and imposed a restriction on the flow, for $\alpha > \alpha_{cr}$ the critical point was of a nodal type, and did not impose any restriction on the flow^{8,15}. As a result there was a continuum of steady-state solutions for a large viscosity: the location of the sonic point was free to vary over a small but finite range of radii. Muchotrzeb⁸ suggested that the flow was unstable for $\alpha > \alpha_{cr}$.

Physical interpretation of the mathematical properties of accretion flow near r_{ms} for $\alpha > \alpha_{cr}$ remains unclear. Suppose that the viscosity is very small and very gradually removes angular momentum from a given ring of matter. The matter will find its new, nearly keplerian orbit, corresponding to its new angular momentum. At $r = r_{ms}$ keplerian orbits have a minimum of angular momentum, and orbits with $r < r_{ms}$ are dynamically unstable. Therefore, a quasistatic disk cannot exist for $r < r_{ms}$, a well-known result. This phenomenon is related to general relativity, a consequence of a strong gravitational field. In the low-viscosity case, matter is pushed across r_{ms} by a small but finite radial pressure gradient. The situation is similar to the Roche lobe overflow.

The amount of angular momentum removed from a ring of matter during one rotation period increases with viscosity. The new keplerian orbit, corresponding to a new angular momentum will be somewhat farther away. Hence the radial velocity increases with viscosity, a well-known phenomenon. As there is a minimum of angular momentum at r_{ms} , a very small change in angular momentum corresponds to a very large change in the radial distance near r_{ms} ; therefore, the radial velocity may become supersonic because of the direct action of viscosity, rather than because of the radial pressure gradient. This seems to be the reason why the flow is different for $\alpha > \alpha_{\text{cr}}$ than for $\alpha < \alpha_{\text{cr}}$.

The flow properties near r_{ms} in a disk accreting onto a black hole cannot have much effect on the observed luminosity, as the amount of accretion energy released near r_{ms} is very small. However, most models of neutron stars have radii smaller than r_{ms} ⁵. In a newtonian case half of the accretion energy is released in the disk, and half in the boundary layer between the disk and the central object¹⁶, provided the object is not spinning near its rotational break-up. In the relativistic case most of the accretion energy is released in the boundary layer⁶, between r_{ms} and the stellar surface. For example, if the neutron star radius is $r_{\text{NS}} = 3r_{\text{Sch}}$, or $r_{\text{NS}} = 1.5r_{\text{Sch}}$, then the fraction of accretion energy released in the boundary layer is 69% and 86%, respectively, and it may well dominate the observed X-ray luminosity. Therefore, the stability of X-ray luminosity of accreting neutron stars may be governed by the stability of disk accretion near r_{ms} .

Many accreting neutron stars are known to be X-ray pulsars, and these are believed to have strong magnetospheres which control the accretion of matter onto the stellar surface¹⁷. Those X-ray binaries that do not have X-ray pulsars in them show no direct evidence of having any magnetospheres. These include galactic bulge sources, X-ray bursters and recently discovered QPOs^{17,18-22}.

The discovery of quasi-periodic oscillations in a number of bright galactic X-ray sources gave rise to many theoretical speculations¹⁸⁻²². In many theories the QPO phenomenon is somehow related to a weak magnetosphere around the accreting neutron star. However, the QPO sources observed so far in a high, as well as in a low X-ray luminosity state, show no evidence for strictly periodic intensity variation, which could be interpreted as the rotation period of a neutron star and its magnetosphere. Even if there is some way to suppress the strictly periodic pulsations, and to hide the hypothetical magnetospheres, it is certainly fair to say that there is no evidence for their existence. At the same time all published models of the QPO phenomenon ignore the unsolved problem of the simplest disk accretion flow through r_{ms} , with a constant α parameter and exact axial symmetry. If that flow cannot be steady for $\alpha > \alpha_{\text{cr}}$, then its instability may offer the simplest explanation of the QPO phenomenon, and possibly for the rapid burster²³.

Preliminary attempts to elucidate the nature of disk accretion flow through r_{ms} , and the sensitivity of results to different formulations of α viscosity, and to different numerical techniques, imply that peculiarities of the flow with $\alpha > \alpha_{\text{cr}}$ are always there, but their consequences are not clear so far (B.P. and M. Rupen, in preparation). A possible relation between the flow through r_{ms} and the QPO phenomenon and the rapid burster remains a matter of speculation. The main purpose of this report is to point out the relevance of the flow through r_{ms} for non-magnetospheric disk accretion onto neutron stars. This relevance is important because the flow properties are governed by the strong gravitational field as described by general relativity.

This work was partly supported by the NSF.

5. Kluźniak, W. & Wagoner, R. V. *Astrophys. J.* **297**, 548-554 (1985).
6. Syunyaev, R. A. & Shakura, N. I. *Pis'ma Astr. Zh.* **12**, 286-292 (1986).
7. Muchotrzeb, B. *Acta Astr.* **33**, 79-87 (1983).
8. Muchotrzeb-Czerny, B. *Acta Astr.* **36**, 1-17 (1986).
9. Novikov, I. & Thorne, K. S. in *Black Holes* (eds DeWitt, B. & DeWitt, C.) 343-450 (Gordon & Breach, New York, 1973).
10. Kozłowski, M., Jaroszyński, M. & Abramowicz, M. A. *Astr. Astrophys.* **63**, 209-220 (1978).
11. Abramowicz, M. A., Jaroszyński, M. & Sikora, M. *Astr. Astrophys.* **63**, 221-224 (1978).
12. Liang, E. P. T. & Thompson, K. A. *Astrophys. J.* **240**, 271-274 (1980).
13. Muchotrzeb, B., and Paczyński, B. *Acta Astr.* **32**, 1-11 (1982).
14. Loska, Z. *Acta Astr.* **32**, 13-24 (1982).
15. Matsumoto, R., Kato, S., Fukue, J. & Okazaki, A. T. *Publ. Astr. Soc. Japan* **36**, 71-85 (1984).
16. Pringle, J. E. A. *Rev. Astr. Astrophys.* **19**, 137-162 (1981).
17. Joss, P. C. & Rappaport, S. A. *Rev. Astr. Astrophys.* **22**, 537-592 (1984).
18. Alpar, M. A. & Shaham, J. *Nature* **316**, 239-241 (1985).
19. Hameary, J. M., King, A. R. & Lasota, J. P. *Nature* **317**, 597-598 (1985).
20. Lamb, F. K., Shibasaki, N., Alpar, M. A. & Shaham, J. *Nature* **317**, 681-687 (1985).
21. Boyle, C. B., Fabian, A. C. & Guilbert, P. W. *Nature* **319**, 648-649 (1986).
22. Alpar, M. A. *Man. Not. R. astr. Soc.* **223**, 469-478 (1986).
23. Lewin, W. G. H. & Joss, P. C. in: *Accretion Driven Stellar X-ray Sources*. (eds Lewin, W. H. G. & van den Heuve, E. P. J.) 41-65 (Cambridge University Press, 1983).

Discovery of an infrared-luminous quasar

J. Patricia Vader* & M. Simon†

* Department of Astronomy, Yale University, Box 6666, New Haven, Connecticut 06511, USA

† Astronomy Program, State University of New York at Stony Brook, Stony Brook, New York 11794, USA

During a redshift survey of an infrared complete sample of galaxies recorded by the Infra-Red Astronomical Satellite (IRAS) we have identified the IRAS source 00275-2859 as a quasar with a redshift of $z = 0.28$. The object looks stellar on both blue and red photographic plates, with an apparent optical magnitude $J = 17.1$ mag. It is a radio source. Its far infrared luminosity is $1.5 \times 10^{12} L_{\odot}$, which exceeds both its optical and radio luminosity. This IRAS source is the second previously unidentified quasar selected through its infrared emission after IRAS source 13349+2438¹. This suggests that the IRAS data contain a new class of quasars with strong internal reddening that possibly share the violent star formation activity of the most luminous IRAS objects.

Optical spectra of the IRAS source 00275-2859 were obtained with the two-dimensional Frutti detector at the 1-m telescope at the Cerro Tololo Inter-American Observatory, as part of a programme to measure redshifts of a 60- μm wavelength flux limited sample of IRAS galaxies^{2,3}. The sample contains 92 IRAS galaxies in seven fields, each approximately $4.5^{\circ} \times 4.5^{\circ}$, selected from the calibrated photographic plates used by

Table 1 Properties of IRAS 00275-2859

Wavelength/band	Source	
12 μm	IRAS	<0.12 Jy (3σ)
25 μm	IRAS	0.23 ± 0.06 Jy
60 μm	IRAS	0.66 ± 0.04 Jy
100 μm	IRAS	0.78 ± 0.07 Jy
J	Schmidt Plate (KOS)	17.1 mag
J-F	Schmidt Plate (KOS)	0.8 mag
6 cm	VLA	20 mJy
20 cm	VLA	50 mJy
Luminosities:		
M_J		-23.0 mag
L (60 μm)		$1.5 \times 10^{12} L_{\odot}$
Emission lines	median cz (km s ⁻¹)	FWHM (km s ⁻¹)
[O II] 3,727 Å	83,925	600
H γ 4,340 Å	83,733	
H β 4,861 Å	83,720	2,700
[O III] 5,007 Å	83,520	800
		Equivalent width (rest frame, Å)
		9
		16
		68
		18

Received 19 February; accepted 23 March 1987.

1. Van der Klis, M. *et al.* *IAU Circ. No.* 4043 (1985).

2. Van der Klis, M. *Nature* **316**, 225-230 (1985).

3. Lewin, W. H. G. & Van Paradijs, J. *Comments Astrophys.* **11**, 127-135 (1986).

4. Lewin, W. H. G. *The Physics of Accretion onto Compact Objects* (eds Mason, K. O., Watson, M. G. & Whites, N. E.) 177-194 (Berlin, 1986).

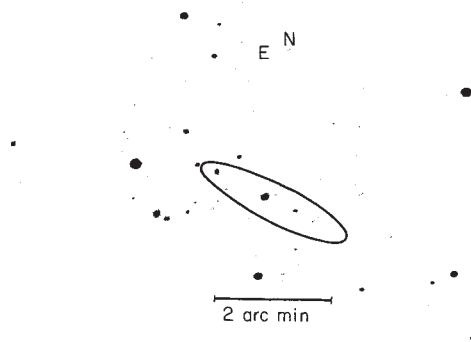


Fig. 1 The optical counterpart of IRAS 00275-2859 is the brightest object in the IRAS position error ellipse. Reproduced from the IIIa-J KOS plate.

Kirshner *et al.*⁴ (KOS) to determine the optical luminosity function of field galaxies. Redshifts are now available for 85% of the objects, which, except for two previously identified Seyfert galaxies and the newly discovered quasar, are galaxies with spectra like that of H II regions.

IRAS 00275-2859 is an unidentified object in the *IRAS Point Source Catalog*⁵, its 60- μm flux density is S_{ν} (60 μm) = 0.66 Jy. It is in the area of KOS field SP3 which is near the south galactic pole. The optical counterpart of IRAS 00275-2859 is by far the brightest object within the IRAS positional error ellipse (Fig. 1). Its optical coordinates and apparent magnitudes, measured by scanning the KOS plates with the Yale Photometric Data Systems (PDS) microdensitometer, are RA(1950) = 00 h 27 min 35.1 s, Dec(1950) = -28° 59' 02", J = 17.1 and $(J - F)$ = 0.8 mag. The optical position is within 12 arc s of the IRAS position.

Two spectra of the object were taken on two different nights in the 4,500-7,000 Å region at 15-Å resolution, the first with an integration time of 5,200 s and the second with an integration time of 9,000 s. Conditions were mostly non-photometric. A weighted average spectrum is shown in Fig. 2. The spectral lines of [O II] at 3,726 Å, Balmer- γ , Balmer- β and [O III] at 5,007 Å, are prominent at the heliocentric redshift z = 0.28. Strong Fe II multiplet emission at the same redshift is also present. The Balmer- β line has a full width at half maximum (FWHM) of 2,700 km s⁻¹, whereas the forbidden O II and O III lines are about 700 km s⁻¹ wide (FWHM). The object is thus spectroscopically similar to Seyfert Type 1 galactic nuclei. At z = 0.28, the distance to the object is 1,190 Mpc for a Hubble constant, H = 75 km s⁻¹ Mpc⁻¹ and declination parameter q = 0.5.

The available information on the optical line and continuum emission, infrared emission in the IRAS bands, and radio continuum emission is summarized in Table 1. The 12, 25, 60 and 100- μm wavelength flux densities are the result of add-scan processing of the IRAS data. The radio wavelength flux densities are from two 8-minute duration scans obtained with the Very Large Array of the National Radio Astronomy Observatory. The far infrared luminosity of IRAS 00275-2859, $1.5 \times 10^{12} L_{\odot}$, calculated using (νS_{ν}) evaluated at 60- μm wavelength as an estimate of its flux, places the object among the most luminous of IRAS sources. Applying neither extinction nor k corrections, its absolute magnitude is M_J = -23.0 mag, which together with the broad Balmer lines and quasi-stellar appearance qualify IRAS 00275-2859 as a quasar. Its far infrared luminosity exceeds its optical luminosity by about an order of magnitude, and both greatly exceed its radio luminosity between 6- and 20-cm wavelengths.

This object differs from the first discovered infrared-band quasar, IRAS 13349+2438¹ in several respects. Its infrared to optical luminosity ratio exceeds that of the latter by an order of magnitude. The spectral index of IRAS 00275+2859 between 25 and 60 μm is typical of that of previously known active

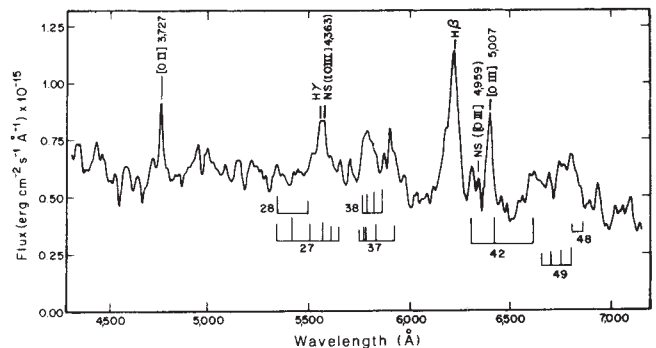


Fig. 2 Weighted average of two flux-calibrated (30% accuracy) optical spectra of 00275-2859. Four emission lines used in the redshift determination and Fe II multiplets are identified. Two forbidden oxygen lines have positions at the redshift of the quasar that coincide with night-sky lines (NS).

galactic nuclei^{6,7,11} in contrast to the unusually flat index of IRAS 13349+2438. The optical colour $(J - F)$ = 0.8 mag of IRAS 00275-2859 is 0.3 mag redder than the $(B - V)$ colour of IRAS 13349+2438 and near the red end of the range observed for quasars⁸. Though radio-quiet, IRAS 00275-2859 is a stronger radio source than IRAS 13349+2438, and has the radio spectral index α = -0.73 ($S_{\nu} \propto \nu^{\alpha}$) characteristic of non-thermal radio sources.

Remarkably, only four of the 242 quasars and active nuclei in the catalogue of Veron-Cetty and Veron⁹ that have been detected in the IRAS survey are optically fainter (m_v > 16.8 mag) than IRAS 00275-2859⁶. The infrared to optical luminosity ratio of the latter quasar is thus unusually large for this class of objects. This supports the suggestion by Rieke¹⁰ and others that the presently known sample of active nuclei may be severely biased against strong infrared emitters, that is, highly obscured dusty nuclei. The discovery of this second infrared luminous quasar substantiates the prediction that the IRAS data may contain a new class of highly obscured active nuclei.

We plan optical imaging of 00275-2859 to elucidate further its nature. If this object follows the trend of decreasing nuclear-fuzz luminosity ratio with increasing $(B - V)$ observed for quasars with detectable fuzz⁸, its red colour suggests a considerable contribution (>50%) to the total optical light by a host galaxy.

We thank Gus Oemler for making the KOS plates available and his PDS and photometry computer programs, Tom Soifer for providing the co-added IRAS data, Carl Bignell, Peggy Perley, and Dave Wunker for obtaining the VLA data, Dick Joyce for obtaining an estimate of the K -magnitude of the object, Charlene Heisler for assistance handling the plate material, Andrew Phillips for help with reduction of the optical spectra, Martin Gaskell for teaching us about Fe II line emission of quasars, and Amos Yahil for help with the Image Reduction and Analysis Facility. J.P.V. is a Visiting Astronomer, Cerro Tololo Observatory/NONO, operated by AURA, Inc., under contract with the NSF.

Received 4 December 1986; accepted 27 February 1987.

1. Beichman, C. A. *et al. Astrophys. J. Lett.* **308**, L1-L5 (1986).
2. Vader, J. P. & Simon, M. in *Star Formation in Galaxies* (ed. Persson, C. J.) (U.S. Government Printing Office, Washington, in the press).
3. Vader, J. P. & Simon, M. *Astr. J.* (submitted).
4. Kirshner, R. P., Oemler, A. & Schechter, P. L. *Astr. J.* **83**, 1549-1563 (1978).
5. *IRAS Point Source Catalog* (US Government Printing Office, Washington) 1549-1563 (1985).
6. Miley, G. K. & de Grijs, R. in *Light on Dark Matter* (ed. Israel, F. P.) 471-486 (Reidel, Dordrecht, 1986).
7. Miley, G. K., Neugebauer, G. & Soifer, B. T. *Astrophys. J. Lett.* **293**, L11-L14 (1985).
8. Hutchings, J. B., Crampton, D. & Campbell, B. *Astrophys. J.* **280**, 41-50 (1984).
9. Veron-Cetty, M. P. & Veron, P. A. *Catalogue of Quasars and Active Nuclei* ESO Scientific Report 1 (1984).
10. Rieke, G. H. *Astrophys. J.* **226**, 550-558 (1978).
11. Neugebauer, G., Miley, G. K., Soifer, B. T. & Clegg, P. E. *Astrophys. J.* **308**, 815-828 (1986).

Variations of stoichiometry and cell symmetry in $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$ with temperature and oxygen pressure

P. Strobel*, J. J. Capponi*, C. Chaillout*,
M. Marezio*† & J. L. Tholence‡

* Laboratoire de Cristallographie CNRS, 166X,
38042 Grenoble Cedex, France

† AT&T Bell Laboratories, Murray Hill, New Jersey 07974, USA

‡ Centre de Recherches sur les Très Basses Températures, CNRS,
166X, 38042 Grenoble Cedex, France

The mixed valence of copper ($\text{Cu(II)}-\text{Cu(III)}$) seems to be important in the recently discovered high-critical-temperature superconductivity of the yttrium barium copper oxide, $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$; several authors have noted the influence of synthesis procedures on the superconducting critical temperature T_c (refs 1, 2). Recent structure determinations³⁻⁵ have shown the presence of oxygen vacancies, changing the copper coordination number from 6 in an ideal stoichiometric perovskite-type compound to 4 and 5 in $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$. Here we report a thermogravimetric study of $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$ carried out at various temperatures and oxygen partial pressures, along with single-crystal X-ray diffraction data showing evidence of a crystallographic transformation as a function of temperature. The optimum copper valence is not reached under usual annealing conditions (800–1,000 °C), even in an oxygen atmosphere, but only below ~350 °C. A change in the rate of oxygen loss with temperature may be related to an orthorhombic to tetragonal structural transformation occurring at ~600 °C in air.

The $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$ phase for the present study was synthesized from stoichiometric proportions of yttrium oxide, barium carbonate and cupric oxide. The reactants were ball-milled twice, fired at 950 °C in air for 12 h, and quenched. Thermogravimetry was carried out in an Uginde-Eyraud thermobalance, using 500- μl platinum crucibles. Heating and cooling rates were in the range 1–2 °C min^{-1} . Atmospheres used were oxygen (Air Liquide, standard grade), air, argon (Air Liquide, N55 grade: <1 p.p.m. O_2) and a 20:1 mixture of argon and oxygen adjusted by high-accuracy flow meters. All experiments were carried out at 1 atm total pressure. The maximum oxygen content for an oxygen pressure of 1 atm is 7 oxygen atoms per formula unit, as shown from a neutron diffraction refinement of the oxygen site occupations on a sample annealed and cooled slowly under flowing oxygen³; this corresponds to a $\text{Cu(III)}/\text{Cu(II)}$ ratio equal to 0.5. Oxygen contents reported by other authors for samples prepared in similar conditions are 6.9 from gravimetric analysis² and 6.81 from another structure refinement from neutron diffraction data⁴.

Figure 1 shows thermogravimetric curves of $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$ samples quenched in air. In both air and oxygen, a mass increase corresponding to ~0.5 oxygen atoms per formula unit occurs in the range 250–350 °C, and a mass loss is observed above 550 °C (in air) or 650 °C (in oxygen). The latter phenomenon is reversible, but not the former: on slow cooling, the samples exhibit no significant variation of mass below ~300 °C (Fig. 1a) and retain their maximum oxygen content ($x=0$).

Such fully oxygenated samples were submitted to thermogravimetry again in oxygen partial pressures ranging from

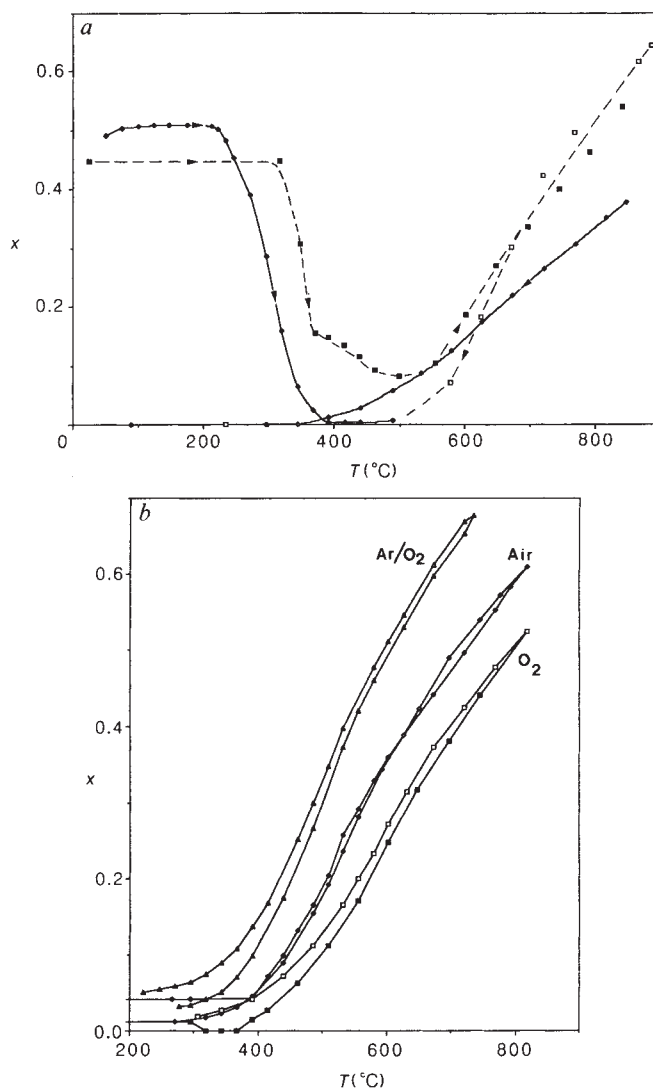


Fig. 1 a, Thermogravimetric curves of $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$ in 1 atm oxygen (continuous line) and in air (dashed line). The data are plotted in terms of the oxygen vacancy content, x , derived from the sample mass by assuming $x=0$ at maximum weight. Arrow-heads indicate heating and cooling trajectories. b, Thermogravimetric curves of $\text{YBa}_2\text{Cu}_3\text{O}_7$ in various atmospheres (closed symbols, heating; open symbols, cooling). The oxygen partial pressure in the argon/oxygen mixture was 0.050 atm.

0.05 to 1 atm (Fig. 1b). All showed a continuous and reversible mass loss starting at ~300 °C. As expected for an oxygen equilibrium, the magnitude of the mass loss increases with decreasing oxygen partial pressure.

Finally, heating in argon (<1 p.p.m. O_2) up to 900 °C resulted in a more pronounced mass loss, corresponding to a partial reduction of Cu(III) and Cu(II) to metallic copper: the crucible was pierced after heating and contained a hard block of a second phase identified as a Pt-Cu alloy.

Several conclusions can be drawn from these experiments. First, the distorted perovskite structure adopted by the Y-Ba-Cu-O can accommodate a considerable range of oxygen vacancies. $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$ samples annealed in air ($x \approx 0.5$) and in oxygen ($x=0$) are both orthorhombic at room temperature. Table 1 lists the cell parameters obtained by a least-squares refinement from X-ray powder data (Guinier camera).

Secondly, the maximum annealing temperature is not the most significant parameter in determining x during the synthesis of $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$. Although high temperatures are obviously necessary for kinetic reasons to allow sufficient diffusion for the

Table 1 Cell parameters (in Å) of $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$ samples quenched in air ($x \approx 0.5$) and annealed in oxygen ($x=0$)

x	a	b	c
~0.5	3.841 (5)	3.889 (4)	11.736 (7)
0	3.828 (3)	3.891 (2)	11.679 (4)

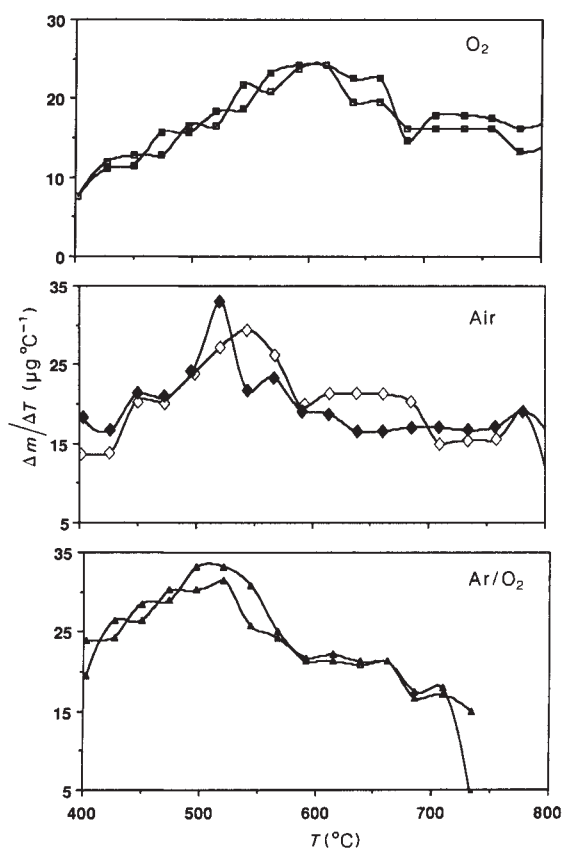


Fig. 2 Data of Fig. 1b plotted as differential mass loss, $\Delta m/\Delta T$, versus T . The maxima in slope at ~ 500 – 600 $^{\circ}\text{C}$ may correspond to a structural transformation (see text).

formation of a single phase, the O₇ stoichiometry is not reached in the 800–1,000 $^{\circ}\text{C}$ annealing stage, even at 1 atm oxygen pressure. The maximum oxidation of copper is achieved only by equilibrating the sample with oxygen below ~ 350 $^{\circ}\text{C}$. Previous authors who obtained the O₇ composition^{2,4} indeed included in their synthesis procedure slow cooling in the furnace atmosphere. The thermodynamics of the Cu(II)–Cu(III) equilibrium are not known. In $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$, Cu(II) and Cu(III) coexist for $x < 0.5$; the range of temperatures and oxygen partial pressures corresponding to $x < 0.5$ in this study is comparable to that in the system La–Ba–Cu–O, previously investigated by Michel and Raveau⁶. Reported losses of superconductivity on annealing in air at 1,100–1,150 $^{\circ}\text{C}$ (ref. 7), or in argon², must correspond to the absence of any Cu(III).

Finally, high-temperature annealing not only prevents the stabilization of Cu(III), but can result in a copper oxidation state lower than two. In the usual annealing range (900–1,100 $^{\circ}\text{C}$), even in an oxygen atmosphere, the system is close to the Cu(II)–Cu(I) boundary: the equilibrium oxygen pressure for the Cu_2O –CuO reaction is ~ 0.1 atm at 1,000 $^{\circ}\text{C}$ and 1 atm at 1,130 $^{\circ}\text{C}$ (calculated from ref. 8), and the thermodynamics of the Cu(I)–Cu(II) couple in ternary or quaternary oxides should not be markedly different⁸.

Our syntheses showed that the materials with highest T_c (onset at 100 K) and sharpest transition (width 5 K) were those annealed in oxygen, then slowly cooled in the furnace, whereas compounds obtained by quenching from high temperature exhibited lower T_c (onset at 65 K), an important decrease in superconducting volume, and less sharp superconducting transitions (width 20 K), as well as reflection broadening in their X-ray patterns (yielding larger standard deviations on the cell parameters; see Table 1). These findings are consistent with the thermogravimetric study: quenching is likely to result in an average oxygen stoichiometry of less than seven atoms per

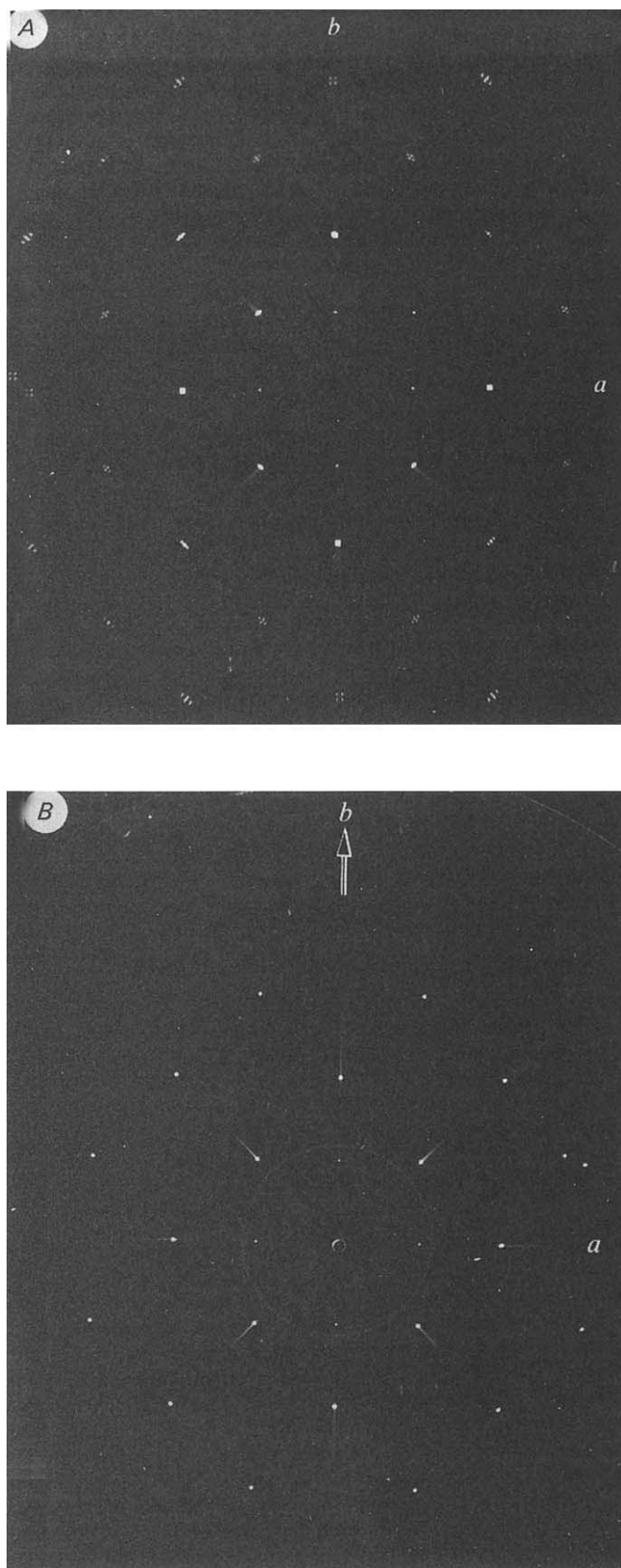


Fig. 3 Precession photographs of a crystal of $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$: A, at room temperature, showing quadruple twinning; B, at 800 $^{\circ}\text{C}$, showing a single tetragonal crystal (rings are due to the high-temperature glue used).

formula unit, and compositional inhomogeneity (variable x) within the sample.

All the curves in Fig. 1b exhibit a maximum in slope at ~ 500 – 600 °C and $x \approx 0.3$. This effect, which can be better seen on derivative thermogravimetric curves (Fig. 2), could be related to a crystallographic transformation. Precession photographs on an oxygen-annealed single crystal indeed show that the structure is tetragonal above ~ 600 °C, whereas twinning is visible on the same crystal at room temperature (Fig. 3). The small distortion of the low-temperature cell with respect to tetragonal symmetry explains why attempts to grow single crystals of $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$ lead to severely twinned crystals. Contrary to previous assumptions⁴, the tetragonal phase is found on the high-temperature side of the transition, corresponding to a significant fraction of oxygen vacancies ($x \geq 0.3$). Work is in progress to investigate the structural arrangements at various temperatures and oxygen partial pressures, and to define quenching conditions for obtaining homogeneous samples.

The Laboratoire de Cristallographie CNRS and the Centre des Recherches sur les Très Basses Températures CNRS are both associated with the Université Scientifique et Médicale de Grenoble.

Received 22 April; accepted 1 May 1987.

1. Wu, M. K. *et al.* *Phys. Rev. Lett.* **58**, 908–910 (1987).
2. Cava, R. J. *et al.* *Phys. Rev. Lett.* **58**, 1676–1679 (1987).
3. Capponi, J. J. *et al.* *Europhys. Lett.* **3**, 1301–1308 (1987).
4. Beno, M. A. *et al.* *Appl. Phys. Lett.* (submitted).
5. David, W. I. F. *et al.* *Nature* (submitted).
6. Michel, C. & Raveau, B. *Rev. Chim. Miner.* **21**, 407–425 (1984).
7. Gallagher, W. J., Sandstrom, R. L., Dinger, T. R., Shaw, T. M. & Chance, D. A. *Solid St. Commun.* (submitted).
8. Kubaschewski, O. & Alcock, C. B. *Metallurgical Thermochemistry*, 5th edn (Pergamon, Oxford, 1979).

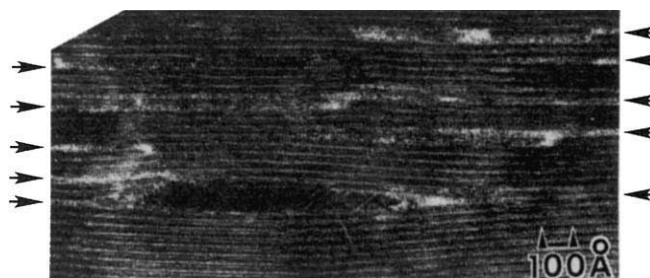


Fig. 1 Low-magnification lattice image of $\text{YBa}_2\text{Cu}_3\text{O}_{6.9}$ in the $[100]$ projection, showing an area of high defect density in this material. The defects are comprised of planar faults and misoriented crystallites, some of which are arrowed. Note the bowing out of the planes near the larger defect. Amorphous regions, not shown in this image, were also observed.

that can be modelled as extrinsic faults. These defects reduce the O/Cu ratio of the material from 2.33 (for the inferred unit cell) towards the experimentally observed value of 2.3 (ref. 2).

Superconducting samples of $\text{YBa}_2\text{Cu}_3\text{O}_{6.9}$, characterized as single-phase by X-ray diffraction, were mechanically polished and ion-milled by 3-keV Ar or Ar/O₂ bombardment to perforation. No differences were observed between samples milled with and without oxygen. By examining the samples at accelerating voltages ranging from 100 to 400 kV, we have ascertained that the features we observe are not due to radiation damage by the electron beam. The samples were examined at the National Facility for Electron Microscopy, Arizona State University, by electron diffraction in the convergent beam (CBED) as well as selected-area modes, and imaged at 5-Å resolution in the temperature range 40–270 K, but no clear evidence of structural changes were observed. Electron channelling experiments using the ALCHEMI technique⁷ were unsuccessful owing to the high defect density (see below). High-resolution lattice images were obtained at AT&T Bell Laboratories, Holmdel, at room temperature, with a JEOL 4000-EX high-resolution transmission electron microscope, capable of a point-to-point resolution of <1.6 Å. Computer-simulated images were produced by a full numerical solution of the Schrödinger equation by the multi-slice method, including all zero Laue zone elastic multiple scattering effects⁸. Comparison of simulated and experimental images over a range of thickness and defocus values then allows us to distinguish between the structural models obtained from different X-ray diffraction experiments^{3,4}.

A macroscopic piece of 'single-phase' $\text{YBa}_2\text{Cu}_3\text{O}_{6.9}$ is polycrystalline, with an average grain size of ~ 0.5 μm. The grains are sometimes separated by an amorphous matrix, which comprises $\sim 10\%$ of the material, and was shown by X-ray microanalysis to be barium-rich compared to the crystalline phase. Our CBED patterns indicate mirror planes normal to both the a - and b -axes, suggesting a space group $Pmmm$ (ref. 3), other workers have proposed $Pm2m$ ^{9,10}. High-resolution lattice imaging shows the grains to come into intimate atomic contact with each other, forming high- and low-angle grain boundaries. Lattice images of grains in the $[100]$ or $[010]$ orientation reveal the presence of planar defects, as well as amorphous and misoriented crystalline inclusions (Fig. 1). (The small difference between the a and b axes prevents us from reliably distinguishing between them; we therefore do not differentiate here between the $[100]$ and $[010]$ directions.) The density of these defects is sometimes so high as to preclude the production of high-quality, simply interpretable lattice images. We have obtained lattice images in the $[100]$, $[110]$ and $[001]$ orientations from areas containing few defects. Figure 2a is a $[100]$ atomic structure image, obtained from a region of crystal sufficiently thin to allow straightforward image interpretation, and Fig. 2b shows a reverse contrast lattice image. Comparison of experimental and simulated images allows us to deduce the metal atom positions shown in Fig. 3a.

Microstructure, oxygen ordering and planar defects in the high- T_c superconductor $\text{YBa}_2\text{Cu}_3\text{O}_{6.9}$

A. Ourmazd*, J. A. Rentschler*, J. C. H. Spence†, M. O'Keeffe†, R. J. Graham†, D. W. Johnson Jr‡ & W. W. Rhodes‡

* AT&T Bell Laboratories, Holmdel, New Jersey 07733, USA

† Department of Physics, Arizona State University, Tempe, Arizona 85287, USA

‡ AT&T Bell Laboratories, Murray Hill, New Jersey 07974, USA

Several of the theories proposed to account for the unusually high critical temperature (T_c) of $\text{YBa}_2\text{Cu}_3\text{O}_{6.9}$ rely on a reduced dimensionality of this system: the stronger electron–phonon coupling in lower-dimensional systems of marginal stability may allow high- T_c superconductivity by the conventional phonon-mediated mechanism¹. X-ray powder diffraction experiments have shown the structure to consist of a tripled perovskite unit cell² which, in the presence of nine oxygen atoms, is three-dimensionally connected. The theoretical proposals therefore depend crucially on the details of the atomic arrangement to reduce the dimensionality of the system. Recent X-ray work on small single-crystal grains^{3,4} has established the structure of the metal framework in this system, but disagreements persist regarding the disposition of the oxygen vacancies needed to achieve the experimentally observed stoichiometry. Moreover, no conclusive information is available regarding the presence of long-range order in the arrangement of the oxygen vacancies, which could give rise to further reduced-dimensional features in this system^{5,6}. Here we report the results of a high-resolution transmission electron microscope study of $\text{YBa}_2\text{Cu}_3\text{O}_{6.9}$, which elucidates the microstructure of this system on an atomic scale. Our results can be consistently interpreted in terms of a particular form of long-range order in the disposition of the oxygen vacancies. We also observe twins, and planar defects

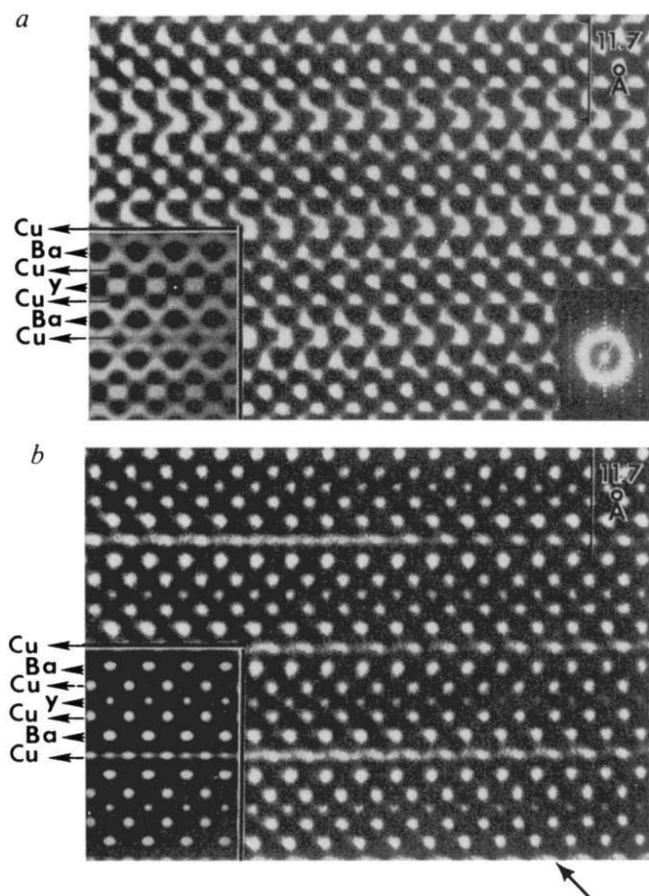


Fig. 2 High-magnification lattice images of the perfect material in the [100] projection. Insets are simulated images and optical diffractogram. *a*, Atomic structure image at optimum (Scherzer) defocus (-420 Å). The metal atom-columns are black. The optical diffractogram confirms the choice of defocus. Sample thickness, 12 Å. *b*, Reverse-contrast image at a defocus of -600 Å. Oblique viewing of this image along the arrow directly reveals the buckling of the Cu planes neighbouring the Y plane. Sample thickness, 55 Å. Simulations show that for this image the electron beam is inclined to the exact [100] direction by ~ 5 mrad.

Because of the small scattering power of the light oxygen atoms, images at very small thicknesses are relatively insensitive to the precise arrangement of these atoms. Dynamical (multiple scattering) effects, which become important with increasing sample thickness, impart significant sensitivity to the exact arrangement of the oxygen atoms in the unit cell¹¹. In general, the interpretation of such effects requires an extremely accurate knowledge of experimental parameters, and thus complicates reliable structure determination. Our simulations show, however, that near optimum defocus (for example, in Fig. 2*a* but not Fig. 2*b*), the strong white features separating neighbouring Ba atom-columns show a dependence on the sample thickness, which is characteristic of a given oxygen arrangement around the basal plane. In particular, for the model consisting of an ordered basal ($z = 0$) chain of O-Cu-O lying perpendicular to the plane of the image, the barium atom positions on either side of the basal plane are separated by a strong white feature, which becomes a continuous white line at thicknesses greater than ~ 60 Å. When the O-Cu-O chains lie in the plane of the image, or when all four possible oxygen sites in the basal Cu plane are fully³ or half² occupied, the strongest white feature appears at the Y planes and does not become continuous with increasing thickness. As we can directly distinguish between the Ba and Y atom-columns at optimum (Scherzer) defocus, this feature constitutes a means of distinguishing between competing models for the oxygen arrangement in the unit cell. We have

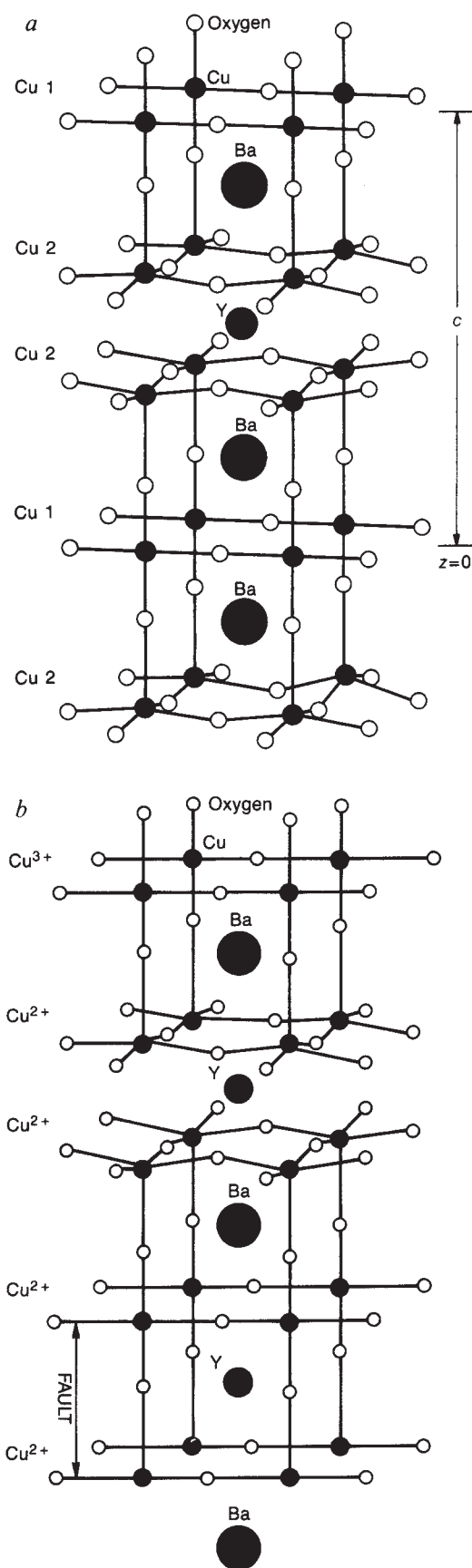


Fig. 3 *a*, Schematic representation of the perfect crystal structure. *b*, Schematic representation of structure containing a planar fault. The valences of the Cu atoms are shown, as determined by charge neutrality considerations.

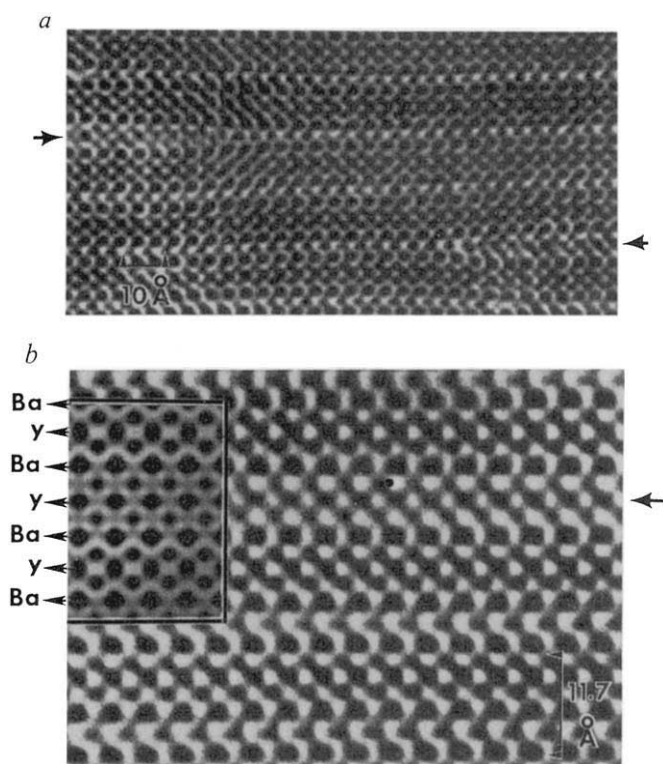


Fig. 4 *a*, Atomic structure image showing two planar faults (arrowed). *b*, High-magnification atomic structure image of region containing a planar fault (arrowed). Inset is the simulated image. Metal atom-columns are black. Sample thickness, 12 Å; defocus, -420 Å.

used the Pendellösung oscillations and image simulation to calibrate the sample thickness. Assuming that the metal atom positions and stoichiometry do not change systematically with thickness, our analysis of the thickness dependence of experimental images suggests that half of the oxygen sites in the basal Cu plane at $z=0$ are empty. The resulting O-Cu-O chains lie perpendicular to plane of Fig. 2, and display long-range order over our field of view (~ 500 Å).

The planar defects in this material are of two types: twin boundaries and extrinsic planar faults. The twins come about as a consequence of the tetragonal to orthorhombic distortion and the accompanying formation of ordered oxygen chains in the basal planes. Thus, a twin boundary causes a 90° rotation in the direction of the O-Cu-O chains. The extrinsic planar faults are shown in Fig. 4. A search of naturally occurring (that is, low-energy) perovskite-related systems, including the K_2NiF_4 and barium titanate structures, suggests that these planar faults have the structure shown in Fig. 3*b*. They consist of a pair of planes similar to the basal planes at $z=0$ separated by a Ba or Y layer; thus, the additional material has stoichiometry $ACuO_2$ (where A is Ba or Y). Comparison of experimental and simulated images suggests that Y is present at the midplane of the defect. This is consistent with Cu^{2+} at the fault planes, as shown in Fig. 3*b*. A system consisting of a periodic repetition of the entire structure shown in Fig. 3*b* has a stoichiometry of $Y_2Ba_2Cu_4O_9$, with an O/Cu ratio of 2.25. The stoichiometry of the perfect structure results in an O/Cu ratio of 2.33. As the experimentally measured O/Cu ratio is 2.30 (± 0.02) (ref. 2), the above two values straddle the experimental measurement. We thus believe that these defects represent the material's response to stoichiometry constraints. However, unlike similar composition defects in other oxides (see, for example, ref. 12), all those we observe terminate within the crystal, indicating a slight misfit with the matrix. Note also that the presence of these defects results in a slight rotation of the planes on either side of the

defect, which is evident from lattice images and their associated optical diffractograms, as well as from transmission electron diffraction patterns of this material. This may result from a 180° twinning of a slightly monoclinically distorted unit cell.

The presence of planar defects in $YBa_2Cu_3O_{6.9}$ must be taken into account in estimates of magnetic susceptibility, and may influence the electronic properties of the material through doping effects. The defects may also act as flux pinning sites, and thus be useful in optimizing extrinsic properties such as the critical current density and magnetic field. Finally, electrons may be localized on these planar faults, and some defects are known to provide strong electron-lattice coupling. But at present we have no means of determining whether defects play any direct role in the superconducting properties of these materials.

We acknowledge valuable discussions with Janet Brown, J. M. Cowley, W. J. Skocpol and S. Zahurak. The low-temperature experiments were partially supported by the NSF.

Note added in proof: We have recently learned of neutron diffraction experiments by J. J. Capponi *et al.* *Europhys. Lett.* **3**, 1301-1308 (1987) which also indicate the presence of long-range order in the oxygen vacancies.

Received 16 April; accepted 6 May 1987.

1. Bardeen, J., Cooper, L. N. & Schrieffer, J. R. *Phys. Rev.* **106**, 62 (1957).
2. Cava, R. J. *et al.* *Phys. Rev. Lett.* **58**, 1676-1679 (1987).
3. Siegrist, T., Sunshine, S., Murphy, D. W., Cava, R. J. & Zahurak, S. M. *Phys. Rev. B* (in the press).
4. Angle, R. *et al.* *Eos* **68**, 161 (1987).
5. Mattheiss, L. F. & Hamann, D. R. *Solid St. Commun.* (submitted).
6. Yu, J., Freeman, A. J. & Xu, J. H. *Phys. Rev. Lett.* **58**, 1035-1037 (1987).
7. Spence, J. C. H. & Taftø, J. J. *Microsc.* **130**, 147-154 (1983).
8. Goodman, P. & Moodie, A. *Acta crystallogr.* **A30**, 280-290 (1974).
9. Beyers, R. *et al.* *Appl. Phys. Lett.* (submitted).
10. Chen, C. H., Werder, D. J., Kwo, J. R., Liou, S. H. & Hong, M. *Phys. Rev. B* (in the press).
11. Anstis, G. R., Lynch, D. F., Moodie, A. F. & O'Keefe, M. A. *Acta crystallogr.* **A29**, 138-147 (1973).
12. Nobel Symposium *Chem. Scripta* **14** (1978-79).

Structure and crystal chemistry of the high- T_c superconductor $YBa_2Cu_3O_{7-x}$

W. I. F. David*, W. T. A. Harrison*, J. M. F. Gunn*, O. Moze†, A. K. Soper*, P. Day†, J. D. Jorgensen‡, D. G. Hinks‡, M. A. Beno‡, L. Soderholm‡, D. W. Capone II‡, I. K. Schuller‡, C. U. Segre§, K. Zhang§ & J. D. Grace||

* Neutron Division, Rutherford Appleton Laboratory, Chilton, Didcot OX11 0QX, UK

† Inorganic Chemistry Laboratory, South Parks Road, Oxford OX1 3QR, UK

‡ Argonne National Laboratory, Argonne, Illinois 60439, USA

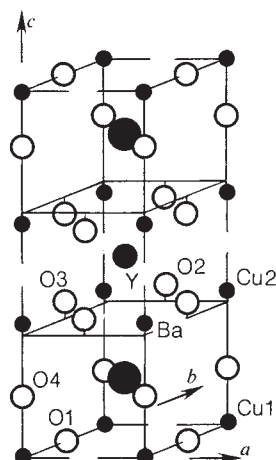
§ Illinois Institute of Technology, Chicago, Illinois 60616, USA

|| Western Michigan University, Kalamazoo, Michigan 49008, USA

Materials within the phase diagram Y_2O_3 -BaO-CuO have recently generated great interest because of the observation of superconductivity with a critical temperature (T_c) of ≥ 90 K^{1,2}; the current consensus is that $YBa_2Cu_3O_{7-x}$ ($x \approx 0.2$) is the superconducting phase (see, for example ref. 3). Hazen *et al.*⁴ concluded, from a single-crystal X-ray diffraction experiment, that its structure is based on the corner-linked octahedral perovskite structure: the tetragonal unit cell (space group $P4m2$) arises from a tripling of the c -axis resulting from the ordering of yttrium and barium cations and associated oxygen defects. Here we present high-resolution neutron powder diffraction data for this phase with $x = 0.15$ (7). Our results agree with ref. 4 with respect to cation positions, but the location of one set of oxygen defects is substantially different and necessitates the lowering of lattice symmetry

† Present address: Istituto MASPEC del CNR, Via Chiavari 18/A, Parma 43100, Italy.

Fig. 1 The structure of $\text{YBa}_2\text{Cu}_3\text{O}_{6.85}$ showing the ordered stacking of yttrium, barium and oxygen defects along the c -axis. Cu(1) is square-planar coordinated; Cu(2) is surrounded by a square pyramid of oxygens.



from tetragonal to orthorhombic (space group $Pmmm$), in agreement with separate findings⁵⁻⁸. In contrast to the 35-K superconductor, $\text{La}_{1.85}\text{Ba}_{0.15}\text{CuO}_4$, the corner-linked CuO_4 planar groups are connected not only as sheets in the a - b plane but also as chains parallel to the b -axis. The average copper valence is 2.23 (5). From bond valence arguments we infer that the Cu^{2+} and Cu^{3+} ions preferentially occupy square pyramidal and square planar sites respectively. Crystal chemical considerations suggest that these structural features may be involved in the superconducting mechanism.

$\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$ was prepared by high-temperature sintering of the appropriate oxides in an oxygen atmosphere. Powder neutron diffraction patterns were obtained at 80 K, 150 K and room temperature on the high-resolution powder diffractometer (HRPD) at the ISIS spallation neutron source, Rutherford Appleton Laboratory⁹. Results from only the room-temperature diffraction data are presented as no detectable change of structure was observed as a function of temperature. Simulation of the structure reported by Hazen *et al.*⁴ showed clear structural similarities with the HRPD diffraction data. Lowering lattice symmetry by introducing a small orthorhombic distortion (space group symmetry reduction from $P4mm$ to $Pmm2$) resulted in excellent agreement between observed and calculated peak positions. The space group symmetry was raised from $Pmm2$ to $Pmmm$ as no evidence for a lack of centre of symmetry was found.

Structural analysis was performed using the Rietveld method based on the Cambridge Crystallography Subroutine Library (CCSL)¹⁰. The cations were located before refinement in the

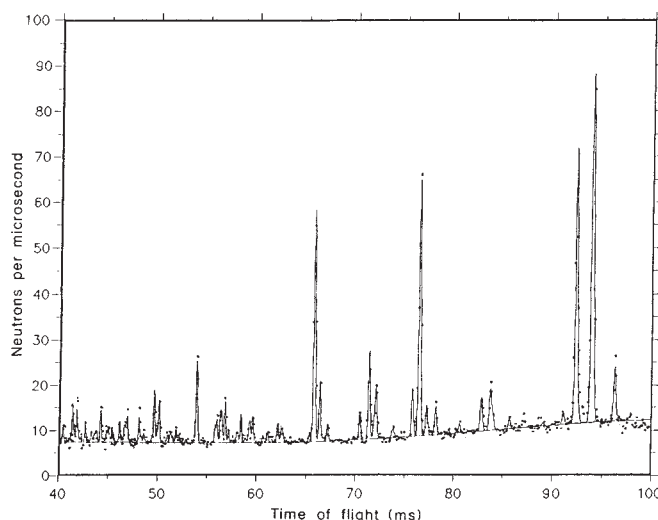


Fig. 2 Final observed (points) and calculated (line) profiles for $\text{YBa}_2\text{Cu}_3\text{O}_{6.85}$ time-of-flight powder neutron data. The d -spacings range from 0.8 to 2 Å in the figure. (A d -spacing of 1 Å corresponds to a time of flight of ~50 ms.)

positions obtained by Hazen *et al.*⁴. The positional parameters of all possible oxygen positions were based on the ideal perovskite structure and refined along with the oxygen site occupancies. Results of the profile refinement (Table 1) indicate two principal sets of oxygen vacancies compared with the ideal perovskite structure: one a - b -plane layer of oxygen ions surrounding the yttrium cation and one line of oxygen ions parallel to the b -axis (see Fig. 1). The location of the latter vacancies are significant, because, despite the pseudo-tetragonal lattice, no structural phase transition to tetragonal symmetry is possible in this material without the disordering of these oxygen vacancies. This is crucial to the understanding of superconductivity in these materials. Detailed analysis also showed that one site (O(2) at $(1/2, 0, 0.3788(5))$) was not fully occupied ($92.5 \pm 3.5\%$ occupancy). A further site (O(1) at $(0, 1/2, 0)$) appeared to be slightly depleted and to possess an anomalously large thermal parameter: with the present quality of data the O(1) site occupancy and temperature factors were highly correlated and thus the O(1) site was constrained to be fully occupied. An anisotropic temperature factor was included in the refinement for O(1). No disordering of the yttrium and barium cations was detected. The refined chemical composition was calculated to be $\text{YBa}_2\text{Cu}_3\text{O}_{6.85(7)}$, corresponding to an average copper valence of 2.23(5) (23.5% Cu^{3+} ; 77.5% Cu^{2+}). The final χ^2 agreement factor $((R_w/P/R_E))^2$ was 0.98 for 28 variable atomic and profile parameters and 302 reflections. Final observed and calculated profile plots are shown in Fig. 2.

The barium ions are tenfold coordinated by oxygen ions that form a cuboctahedron with two vertices missing. Yttrium is eightfold coordinated by an approximate cube of oxygen ions. Bond distances and angles for these polyhedra are typical of the species concerned. The copper ions sit in two crystallographically distinct and chemically dissimilar sites. One site (Cu(1) at $(0, 0, 0)$) is surrounded by a square planar oxygen configuration; the second site (Cu(2) at $(0, 0, 0.3555(3))$) is fivefold coordinated by a square pyramidal arrangement of oxygens. Table 2 lists selected bond lengths and angles for the room-temperature structure. The unexpectedly short Cu(1)-O(4) bond length of 1.843 Å, similar to that found in alkali metal cuprates such as KCuO_2 (ref. 11), prompted a bond strength calculation (Table 2) using the method of Brown and Wu¹². This suggests that the Cu^{3+} ions preferentially occupy the square planar Cu(1) site at $(0, 0, 0)$, a typical feature of d^8 ions. Indeed, locating Cu^{2+} and Cu^{3+} ions on square pyramidal, Cu(2), and square planar, Cu(1), sites respectively yields the stoichiometry $\text{YBa}_2\text{Cu}_3\text{O}_7$. Given

Table 1 Final crystallographic data for $\text{YBa}_2\text{Cu}_3\text{O}_{6.85}$ at room temperature

Atom	Wyckoff symbol	x/a	y/b	z/c	$B^* (\text{\AA}^2)$	Site
Y	1h	1/2	1/2	1/2	0.9(1)	
Ba	2t	1/2	1/2	0.1844(5)	0.8(1)	
Cu(1)	1a	0	0	0	0.9(1)	
Cu(2)	2q	0	0	0.3554(3)	0.6(1)	
O(1)	1e	0	1/2	0	†	
O(2)	2s	1/2	0	0.3788(5)	0.3(4)	0.92(3)
O(3)	2r	0	1/2	0.3771(5)	0.6(1)	
O(4)	2q	0	0	0.1579(3)	0.8(1)	

Orthorhombic: space group $Pmmm$ (no. 47); $a = 3.8187(2)$, $b = 3.8833(2)$, $c = 11.6687(6)$ Å.

* Isotropic temperature factor.

† For O(1), 3 anisotropic temperature factors were refined: $B_{11} = 4.8(6)$ Å², $B_{22} = 1.4(4)$ Å², $B_{33} = 2.5(6)$ Å²

$R_N = \sum |I_k(\text{obs}) - I_k(\text{calc})| / \sum I_k(\text{obs}) = 3.9\%$

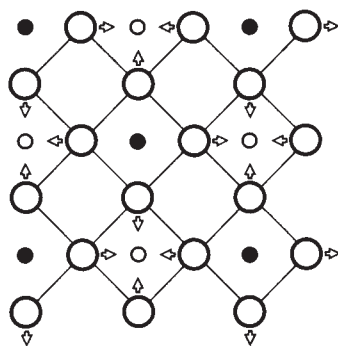
$R_P = \sum |y_i(\text{obs}) - y_i(\text{calc})| / \sum y_i(\text{obs}) = 10.9\%$

$R_{wP} = \{ \sum w_i [y_i(\text{obs}) - y_i(\text{calc})]^2 / \sum w_i y_i^2(\text{obs}) \}^{1/2} = 11.7\%$

$R_E = \{ [N - P + C] / \sum w_i y_i^2(\text{obs}) \}^{1/2} = 11.8\%$

where N , P and C are the number of observations, parameters and constraints respectively.

Fig. 3 Schematic view down the z -axis of the proposed in-plane breathing mode of the copper-oxygen sheets, leading to dynamic localization of electron pairs (Cu^{1+} character, closed circles) and electron vacancies (Cu^{3+} character, open circles). Large circles are oxygens.



the strong inference from bond strength calculations of an average valence of two on the square pyramidal site, the refined stoichiometry, $\text{YBa}_2\text{Cu}_3\text{O}_{6.85(7)}$, gives a 70(15)% : 30(15)% $\text{Cu}^{3+}:\text{Cu}^{2+}$ distribution on the square planar site. These conclusions are at variance with a similar analysis by Capponi *et al.*⁷, probably because of an incorrect choice of the unit strength distance parameter, D1, at 1.74 Å by these authors.

Although the structure of $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$ is substantially different from the 35-K superconductors $\text{La}_{1.85}\text{A}_{0.15}\text{CuO}_4$ ($\text{A} = \text{Ba}, \text{Sr}$), both structures may be considered to be derived from the perovskite structure and to contain corner-linked CuO_4 square planar arrangements. In $\text{La}_{1.85}\text{A}_{0.15}\text{CuO}_4$ these planar groups are two-dimensionally connected in the a - b plane and are derived from a Jahn-Teller-like elongated octahedral (4+2) coordination as in the parent La_2CuO_4 phase. In $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$ two types of planar groups exist: two thirds of the a - b planes containing copper consist of CuO_4 planar-like groups corner-linked in the a - b plane, derived from square pyramidal (4+1) coordination of the $\text{Cu}(2)$ site at (0, 0, 0.3555(3)). These two-dimensional sheets are slightly puckered, with the copper displaced from the oxygen plane by ~ 0.262 Å towards the apical oxygen. The planar copper-oxygen bonds all lie between 1.93 and 1.96 Å. In addition, there are one-dimensional chains of true CuO_4 squashed planar groups oriented in the b - c plane and linked parallel to the b -axis; these are weakly bound to the a - b -plane CuO_4 planar groups by the long (2.308 Å) apical Cu-O bond of the square pyramid. There are no linkages between these square planar groupings in the a direction (see Fig. 1); this is reflected in the large anisotropic temperature factors for O(1).

It is likely that the sheets of corner-linked CuO_4 planar groups are involved in the superconducting mechanism because of their common occurrence in the $\text{La}_{2-x}\text{Ba}_x\text{CuO}_4$ and $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$ structures. An attractive chemical explanation for the occurrence of Cooper pairs in these sheets is the disproportionation of (d^9) Cu^{2+} into (d^{10}) Cu^+ and (d^8) Cu^{3+} , by analogy with the disproportionation of Au^+ and Au^{3+} in $\text{Cs}_2(\text{Au}^+\text{Au}^{3+})\text{Cl}_6$ (ref. 13). This leads to an electron configuration of ($d_{x^2-y^2}$)⁰ and ($d_{x^2-y^2}$)² on alternate copper sites, which is equivalent to a dynamic localization of electron pairs. The dynamic emptying and filling of the ($d_{x^2-y^2}$) orbitals will isotropically contract and stretch the in-plane Cu-O bonds. The breathing mode associated with this effect, shown in Fig. 3, may be expected to behave anomalously around the superconducting transition temperature.

The observation of large temperature factors for O(1) suggests that the chains of CuO_4 planar groups parallel to the b -axis may also be associated with the superconducting mechanism. (These temperature factors may be reliably considered to be large, as examination of the variance-covariance matrix shows no substantial correlation between these and other variables. The temperature factors are similarly large when the O(1) site occupancy is also refined.) The large values of B_{11} and B_{33} indicate large low-frequency librational movements of the CuO_4 groups, but these modes are unlikely to be strongly coupled to the electron motion. The large value of B_{22} , indicative of a stretching and contraction of the Cu-O(1) bond, is more interest-

Table 2 Selected bond distance, angle and strength data for $\text{YBa}_2\text{Cu}_3\text{O}_{6.85}$

Bond	Distance (Å)	Number	Bond	Distance (Å)	Number
Y-O(3)	2.387	× 4	Cu(1)-O(4)	1.843	× 2
Y-O(2)	2.402	× 4	Cu(1)-O(1)	1.942	× 2
Ba-O(4)	2.741	× 4	Cu(2)-O(2)	1.929	× 2
Ba-O(1)	2.877	× 2	Cu(2)-O(3)	1.958	× 2
Ba-O(3)	2.949	× 2	Cu(2)-O(4)	2.306	× 1
Ba-O(2)	2.986	× 2			

Bond	Angle (deg)
O(1)-Cu(1)-O(1)	180
O(1)-Cu(1)-O(4)	90
O(2)-Cu(2)-O(3)	88.96
O(2)-Cu(2)-O(4)	98.14
O(3)-Cu(2)-O(4)	97.42

Bond strength data:					
Assuming:	Cu ²⁺	Cu ³⁺			
Cu(1)-O(4)	2 × 0.33	2 × 0.38			
Cu(1)-O(1)	2 × 0.24	2 × 0.26			
Summed strengths	2.27	2.56	≡ calculated	Cu(1)	
			valence		
Cu(2)-O(2) × 2	2 × 0.23	2 × 0.25			
Cu(2)-O(3) × 2	2 × 0.23	2 × 0.25			
Cu(2)-O(4) × 1	0.17	0.16			
Summed strengths	2.01	2.16	≡ calculated	Cu(2)	
			valence		

Bond strengths, s , are calculated from bond lengths, R , using the formula, $s = (R_0/R)^N$, where R_0 is the bond length which corresponds to unit valence. For Cu^{2+} $R_0 = 1.718$, $N = 6.0$; for Cu^{3+} $R_0 = 1.771$, $N = 7.5$. Cu^{2+} values from Brown and Wu¹⁰; Cu^{3+} values derived from other Cu(III) salts by us. The average 'bond strength' valences of the copper sites may be used to provide an independent rough estimate of the chemical composition as $\text{YBa}_2\text{Cu}_3\text{O}_{6.64}$ (derived from Cu^{2+} parameters) and $\text{YBa}_2\text{Cu}_3\text{O}_{6.94}$ (derived from Cu^{3+} parameters): although reliance should not be placed on this formula, it is in broad agreement with the refined stoichiometry

ing as this mode has the correct symmetry to couple strongly the the electron motion in an analogous fashion to the breathing mode discussed above. Thus electron pairs may also be localized within these chains of CuO_4 planar groups. Another indication of the importance of these chains in the superconductivity is the fact that the bond lengths of the Cu-O bonds in the sheets and along the chains are respectively longer and shorter than in the lanthanum compounds. This implies that the copper electronic bands of the sheets should be more insulating in the yttrium case than the lanthanum case. Moreover the estimate of 30% Cu^{2+} along the chains implies no such tendency to insulating behaviour there. The existence of the chains in $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$ may therefore be the crucial structural feature in determining its high T_c .

We thank the Rutherford Appleton Laboratory for provision of neutron scattering facilities. Part of this work was supported by the US Department of Energy, BES-Materials Sciences.

Received 28 April; accepted 11 May 1987.

- Wu, M. K. *et al. Phys. Rev. Lett.* **58**, 908-911 (1987).
- Hinks, D. G. *et al. Appl. Phys. Lett.* (in the press).
- Rao, C. N. R., Ganguly, P., Raychaudhuri, A. K., Mohan Ram, R. A. & Sreedhar, K. *Nature* **326**, 856-857 (1987).
- Hazen, R. M. *et al. Phys. Rev. Lett.* (in the press).
- Beno, M. A. *et al. Appl. Phys. Lett.* (in the press).
- Greedan, J. E., O'Reilly, A. & Strager, C. V. *Phys. Rev.* (submitted).
- Capponi, J. J. *et al. Europhys. Lett.* **3**, 301-308 (1987).
- Beech, F., Miraglia, S., Santoro, A. & Roth, R. S. *Phys. Rev. Lett.* (submitted).
- Johnson, M. W. & David, W. I. F. *Rutherford Appleton Lab. Rep. No.* 85/112 (1985).
- Matthewman, J. C., Thompson, P. & Brown, P. J. *J. appl. Crystallogr.* **15**, 167-173 (1982).
- Hestermann, K. & Hoppe, R. Z. *anorg. allg. Chem.* **367**, 249-253 (1969).
- Brown, I. D. & Wu, K. K. *Acta crystallogr. B32*, 1957-1959 (1976).
- Tindemans, J. C. M. & Verschoor, G. C. *Mater. Res. Bull.* **9**, 1947-1952 (1974).

Physical and chemical modification of the surface of Venus by windblown particles

Ronald Greeley*, John R. Marshall* & James B. Pollack†

* Department of Geology, Arizona State University, Tempe, Arizona 85287, USA

† NASA-Ames Research Center, Moffett Field, California 94035, USA

Several investigators have speculated on the possibility of wind erosion on Venus¹⁻³. Although winds of sufficient strength to move particles are documented^{4,5} and many investigators suggest active aeolian processes⁶, the style and efficiency of erosion are open to question. Here we present the results of simulations of the venusian surface environment involving windblown grains, which show that significant chemical and physical changes may occur, even in the slow-moving winds recorded on Venus. The edges of grains become worn and shed comminuted debris, which collects on weathered surfaces and grains alike. The resulting transfer of material from loose grains to bedrock surfaces (and vice versa) could yield misleading results on rock composition; moreover, the generation of comminuted debris would enhance chemical reactions that could affect the composition of the atmosphere. These results are thus relevant in assessing rates of surface degradation, the evolution of small-scale (less than a few metres) surface features as seen in images returned from the Soviet Venera missions^{8,9} and in the interpretation of compositional data for surface materials¹⁰.

The Venus Simulator⁷ (Fig. 1) propels sand- or pebble-size particles at controlled speeds and periodicity against rock targets in a carbon dioxide atmosphere at temperatures up to 800 K and pressures up to 114 bars. These conditions are achieved in a pressure vessel which has an internal volume of 0.05 m³. The vessel contains a 0.07-m diameter, 0.75-m-long, tubular furnace which provides an electrically-heated gas reservoir. A grain-propelling device is inserted through the centre of one of the end flanges into the reservoir and is viewed through a 0.05-m-thick quartz window. Illumination of the device is through a light pipe at the opposite end of the pressure vessel. A gas-pulsing system produces rapid-cycle release of internal pressure, and in so doing causes gas to be drawn from the reservoir through a 0.02-m-long gas gun contained within the device. This flow propels particles at a rock target situated directly in the gas flow. The flowing gas is at the same temperature as the impacting grains and the target. Pressure in the vessel is maintained by a gas-intensifier system.

A series of tests was run (Table 1) in which basalt and olivine grains 3 mm in diameter were blown against glassy basalt targets at velocities of 0.5–0.7 m s⁻¹ under atmospheric pressures ranging from 20 to 62 bars. The temperature was maintained at 737 K for all tests except run no. 008. The pressures tested are appropriate for particle motion by gentle venusian winds at elevations ranging from the mid-level plains to the summits of mountains on Venus¹¹. Experiments were also run using the same materials at near-terrestrial conditions for comparison. The surfaces of

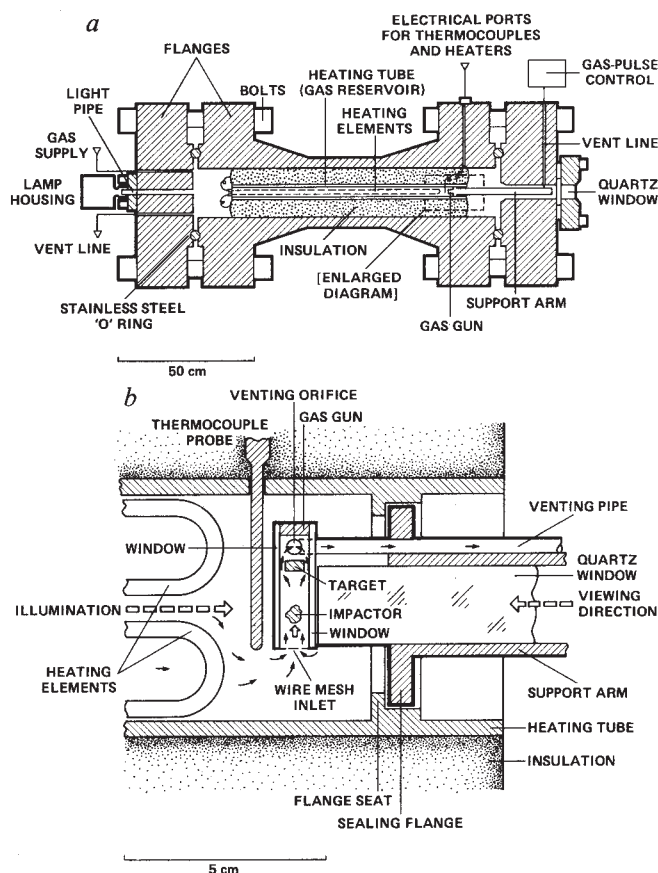


Fig. 1 a, Diagram showing the internal configuration of the Venus Simulator. Target and incident grains are contained within area shown in (b).

the colliding grains and the rock targets were analysed using a scanning electron microscope (SEM) after each run. A Kevex multi-channel X-ray analyser attached to the SEM was used to determine chemical changes on the surfaces of the grains and the targets. The surfaces of the targets were also assessed with a Tencor Alpha-step 200 profilometer. This device traverses the surface with a diamond stylus and converts vertical stylus motion into an electronic signal used to derive a topographic profile with a vertical resolution of 5 nm.

Each run consisted of propelling a single grain against the target face for 10⁵ collisions at a controlled velocity. Target rock samples were placed in the chamber to receive collisions normal to their surfaces; part of the surface was covered for protection against impact and served as a 'control' surface. The periodicity of impact was about two per second. If a single particle on Venus were to saltate with the same periodicity, the experiments would represent approximately 14 hours of particle entrainment in a constantly blowing wind. Transport distance would be ~3 km if the grain saltated ~3 cm in a perfectly repetitive fashion (3 cm is reasonable for grains of the size tested). Ordinarily, winds carry large interacting populations of grains, and neither the wind, nor particle motion, would be so constant. The extrapolation to natural conditions, therefore, is very approximate, but it is probably reasonable in terms of the expenditure of particle kinetic energy per unit time and distance.

Figure 2 shows results from the experiments. Despite the very low impact velocities, the edges of all the colliding grains became rounded. Material was removed by chipping to produce pits, and by a finer-scale abrasion process that led to smoothing of areas between the pits. Average attrition was estimated from analysis of the SEM images as less than 1% volume reduction of the grains. Within the range of conditions tested, the abrasion rate apparently was unaffected by variations in temperature or

Table 1 Preliminary experiments in the Venus Simulator

Run no.	Target	Type of impact grain	Temperature (K)	Pressure (bar)	No. of impacts	Impact velocity (ms ⁻¹)
004	Basalt	Basalt	737	20	10 ⁵	0.7
005	Basalt	Olivine	737	20	10 ⁵	0.7
006	Basalt	Basalt	737	40	10 ⁵	0.5
007	Basalt	Basalt	737	62	10 ⁵	0.5
008	Basalt	Basalt	290	6	10 ⁵	0.6

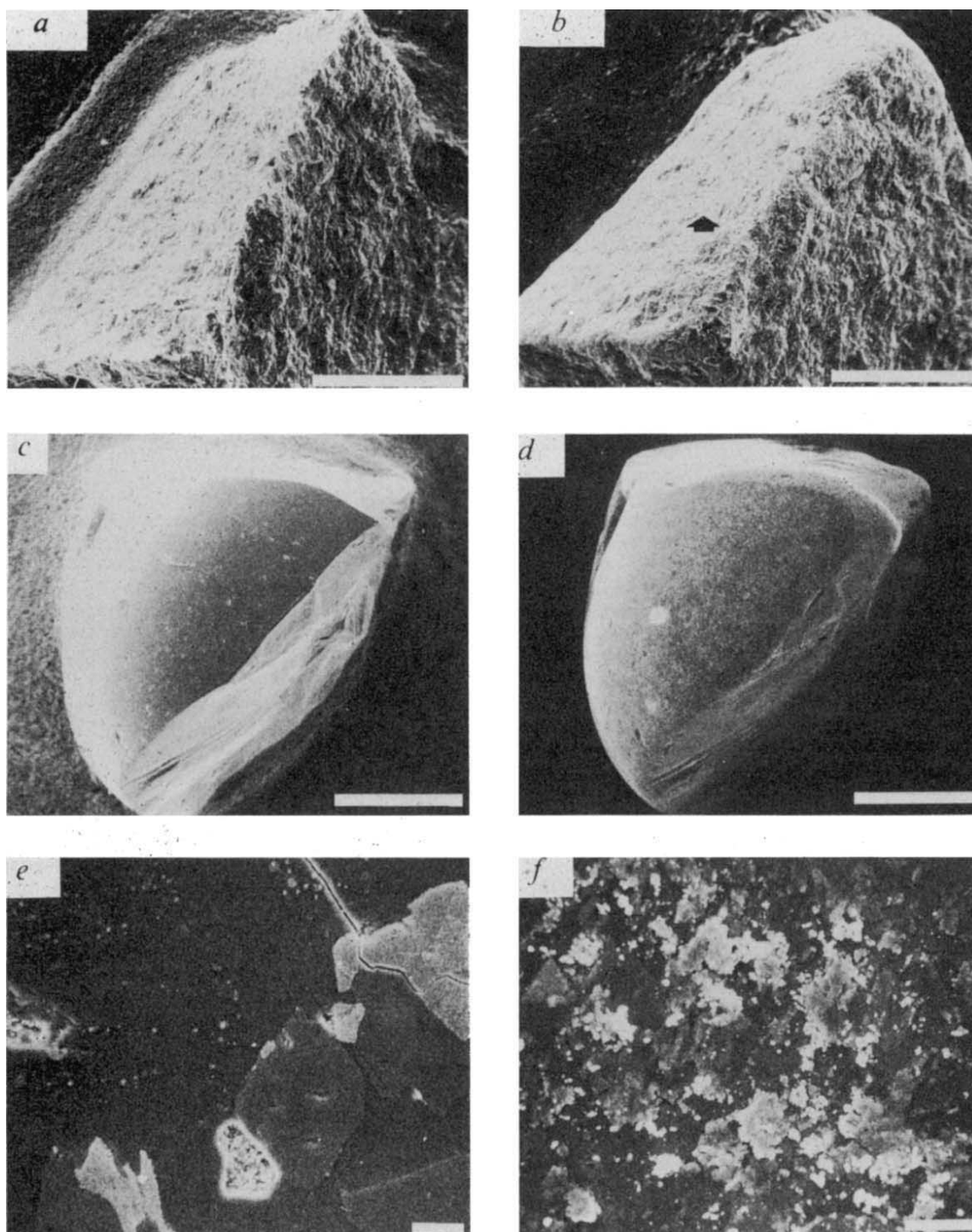


Fig. 2 SEM views of incident grains and basalt target surfaces: *a*, 3-mm basalt grain before collision; *b*, same grain after experiment; arrow indicates erosional pit; *c*, 3-mm olivine grain before experiment; *d*, same olivine grain after impact of grains, sharp edges of grains become rounded and subdued; *e*, surface control area (not weathered by impact of grains) on basalt target; *f*, other half of same target after impact, showing accretion of fine debris derived from impacting grains.

pressure. Moreover, there were no significant differences noted between abrasion of the olivine and basalt grains.

The basalt targets hit by the grains showed no apparent erosion in the experiments, although the surfaces were noticeably modified (Fig. 2). Profiles of the targets before and after each test show an increase in elevation of the weathered surface of several micrometers. Both the surface appearance and the increase in elevation are attributed to accretion of comminuted material derived from the incident grains. Use of the olivine grain in run no. 005 enabled verification of accretion on the

target of material derived from the incident grain. Chemical analysis showed Si, Al and Ca for the basalt target control section and Si, Mg and Fe for the olivine incident grain. The weathered area of the basalt, however, had fine comminution debris on the surface (Fig. 2*f*) and X-rays registered Si, Al, Ca, Mg and Fe. Thus, we infer that Mg and Fe were transferred to the target from the olivine grain. The olivine grain also had fine debris on its surface, but showed no trace of Al or Ca. Debris on the grain was probably 'retrieved' from the accretionary layer.

Run no. 008 was a low-temperature control experiment.

Although the test was run to simulate (in part) an Earth environment, an elevated atmospheric pressure (6 bars) was required to move the grains in the experiment apparatus. The results were markedly different from the high-temperature runs; although abrasion of the incident grain edges occurred, the accretionary layer was absent from the target face. Thus, the process of accretion appears to require the elevated temperatures and atmospheric pressures attained on Venus.

Unknown factors in the experiments are the physical and chemical changes that might occur as a consequence of placing the samples in the high temperature/pressure environment after having reached some stability in the terrestrial environment. At this time, such potential changes are considered to be minimal, as the 'control' surfaces remained unaffected.

Although the experimental results described here are limited in the range of conditions and materials tested, preliminary analyses have implications for several aspects of venusian lithosphere-atmosphere interactions. Perhaps most important is the potential effect on surface compositions. Rocks and outcrops exposed to the impact of windblown grains on Venus may be veneered with a layer of comminuted debris derived from the grains. Depending upon factors such as orientation with respect to wind direction, local turbulence (as around rocks), and duration of exposure, veneers may range from zero to tens of micrometers in thickness. Thus, remote sensing and instruments that obtain *in situ* measurements of composition for only the surface may not be assessing 'bedrock' materials. Although planetologists have long been aware of similar problems arising from mantles of windblown dust settled from the atmosphere¹², the mechanics described here suggest that material can be accreted on vertical surfaces. Thus, even measurements from surfaces that might be expected to be free of dust mantles have no guarantee against compositional mixing.

The experiments described here did not show abrasion of the target surface by the incident grains. However, the current design for the rock target placement in the chamber exposes only a flat surface to impact. It is likely that abrasion of target edges would occur at a rate similar to the abrasion of edges on the impacting grains, if they were exposed. Thus, angular projections on rocks and bedrock surfaces on Venus would probably be abraded, even in gentle winds, as suggested by Garvin *et al.*¹³.

Experimental results also have implications for the chemical weathering regime on Venus. The generation of comminuted debris from windblown grains substantially increases the surface area of material for exposure to the atmosphere. Thus, chemical weathering processes as described by many investigators^{10,14} would be accelerated. Moreover, Nozette and Lewis¹⁵ proposed that windblown grains would effectively buffer the composition of the atmosphere; the generation of finely comminuted material by impact would enhance this process. On the other hand, bedrock surfaces that have received a veneer by the process described here may be partly shielded from abrasion by a more energetic impact of windblown grains and be partly protected from chemical weathering. These aspects will be assessed in future experiments.

Received 13 February; accepted 23 March 1987.

1. Ronca, L. B. & Green, R. R. *Astrophys. Space Sci.* **8**, 59–65 (1970).
2. Sagan, C. J. *Atmos. Sci.* **32**, 1079–1083 (1975).
3. Iversen, J. D., Greeley, R. & Pollack, J. B. *J. Atmos. Sci.* **33**, 2425–2429 (1976).
4. Antsibor, N. M. *et al.* translated from *Kosmicheskie Issledovaniya* **14**, 714–721 (1976).
5. Avduvskii, V. S. *et al.* translated from *Kosmicheskie Issledovaniya* **14**, 710–713 (1976).
6. Greeley, R. *et al.* *Icarus* **57**, 112–124 (1984).
7. Marshall, J. R., Greeley, R., Tucker, D. W. & Pollack, J. B. *Repts. Planet. Geol. Program* (in the press).
8. Florensky, C. P., Ronca, L. B. & Basilevsky, A. T. *Science* **196**, 869–871 (1977).
9. Basilevsky, A. T. *et al.* *Geol. Soc. Am. Bull.* **96**, 137–144 (1985).
10. Volkov, V. P., Zolotov, M. Yu & Khodakovskiy, I. L. in *Advances in Physical Geochemistry* (ed. Saksena, S. K.) 136–190 (Springer, Berlin, 1986).
11. Seiff, A. in *Venus* (eds Hunten, D. M., Colin, L., Donahue, T. M. & Moroz, V. I.) 215–279 (University of Arizona Press, Tucson, 1983).
12. Garvin, J. B. *Proc. Lunar planet. Sci.* **12B**, 1493–1505 (1981).
13. Garvin, J. B., Mouginitz-Mark, P. J. & Head, J. W. *The Moon and Planets* **24**, 355–387 (1981).
14. Florenskii, K. P. *et al.* translated from *Kosmicheskie Issledovaniya* **21**, 351–354 (1983).
15. Nozette, S. & Lewis, J. S. *Science* **216**, 181–183 (1982).

Dating fault gouges

M. Kralik*, K. Klima† & G. Riedmüller†

* Geotechnisches Institut, Bundesversuchsanstalt Arsenal, A-1030 Wien, Austria

† Institut für Technische Geologie, Technische Universität Graz, A-8010 Graz, Austria

Geological faults are fractures in the Earth's crust along which there has been some displacement, and are thus zones of weakness which are of importance in planning civil engineering projects, constructing underground caverns and conducting earthquake research. To assess the potential for future displacements and fluid circulation along faults it is helpful to establish the age of the latest episode of mineral growth in fault zones; however, there have been few attempts to date recrystallized and newly grown minerals developed in fault zones at low temperature^{1–4}. Here we report K–Ar and Rb–Sr ages for clay fractions from two fault zones in the Eastern Alps, and compare them with ages of the same fine fractions from the undeformed host rock. The results suggest that the bulk of the clays along these faults were formed during early Cretaceous and Tertiary times, respectively. No major displacements seem to have taken place in recent times.

The history of low-temperature fillings of faults begins with progressive cracking and consequent reduction of grain size of the host rock and culminates in the development of fault gouges which often consist of crushed rock fragments with varied degrees of replacement and of newly formed minerals, like clays, occurring as filling deposits in the cracks. The composition of the clay minerals is controlled mainly by the chemistry and the amount of circulating fluids^{5–7}. These clay minerals are amenable to isotope analyses to determine the age of deformation.

The investigated Upper Triassic limestone fault (Pailwand, south-east Abtenau) is one of several major, some kilometres-long, north-west to south-east striking fault zones in the southern part of the Northern Calcareous Alps. The fault zone forms a several-metre-wide shear zone with sericitic coatings on the surface of 1–2 centimetre-long tectonic lenses in the Hallstatt Limestone. The fault zone cuts through a tectonic block which was transported into its present position by late Jurassic gravity gliding from elevated areas in the south⁸. The anchimetamorphic overprint in this region is supported by illite crystallinity data and the presence of such metamorphic index minerals as pyrophyllite, paragonite and chloritoid. K–Ar and Rb–Sr isotope data indicate an early Cretaceous age (fine fractions) of this very-low-grade metamorphism⁹.

Fracture zones one-half to several metres wide and filled with cataclastic fault gouges form steep faults running NW–SE and NE–SW and cutting through the area of garnet mica schists of the Kreuzeck Group, which has undergone a complex polymetamorphic history. The two black and green soft fault gouge samples were both collected from a water diversion tunnel (Drassnitzbach Überleitungsstollen). The Kreuzeck Group is part of the Altkristallin basement which was thrust upon the Penninic units exposed in the tectonic Tauern Window 20 km to the north. Detailed K–Ar and Rb–Sr model ages revealed Hercynian ages (muscovite, biotite) in the south and, after a transition zone, Cretaceous ages in the north of the Kreuzeck Group¹⁰. The fault gouge samples are from the transition zone.

The simultaneous tectonic deformation and restructuring or new mineral formation is critical in establishing a meaningful age. Similar effects of restructuring and neoformation of minerals from regional or contact-metamorphism, which frequently occurred in these areas, may obscure the coeval relationship between the neoformation of minerals and the cataclastic deformation. To overcome these possible difficulties, (1) two very fine fractions (μm range, less than 25% overlap in

Table 1 Rb-Sr and K-Ar isotope data of fine fractions in and adjacent to the fault zones

Sample grain size (μm)	Rb ($\mu\text{g g}^{-1}$)	Sr ($\mu\text{g g}^{-1}$)	$^{87}\text{Rb}/^{86}\text{Sr}^*$	$^{87}\text{Sr}/^{86}\text{Sr}^*$ (1σ)	K (%)	$^{40}\text{Ar}^*$ (nl) STP	Rb-Sr model age	Initial ratio	K-Ar model age
Pailwand fault (Halstatt Limestone)									
3L Amb	1.8	243	0.021	0.70838 ± 14	—	—	—	—	—
3 <5	330	18	54.88	0.80516 ± 68	5.59	23.36	96	0.7301	104 ± 11
3 5-12	396	9	130.1	0.90967 ± 66	4.79	20.04	97	0.7301	104 ± 10
3L <5.3 M HCl	13.7	3.4	12.01	0.74708 ± 42	—	—	—	—	—
4L Amb	0.9	253	0.010	0.70868 ± 27	—	—	—	—	—
4 <5	301	21	42.22	0.78992 ± 64	4.90	19.45	100	0.7301	99 ± 13
4 5-12	304	9	100.55	0.87835 ± 99	3.82	15.42	104	0.7301	101 ± 11
4L <5.3 M HCl	11.1	5.5	6.12	0.73842 ± 31	—	—	—	—	—
5L Amb	1.2	178	0.019	0.70792 ± 23	—	—	—	—	—
5 5	348	50	20.48	0.75523 ± 59	6.17	38.75	152	0.7111	155 ± 15
5 5-12	343	13	79.37	0.86655 ± 99	4.14	23.24	138	0.7111	139 ± 9
5L <5.3 M HCl	9.7	17.2	1.642	0.71153 ± 44	—	—	—	—	—
Kreuzeck faults (Garnet mica schist)									
262 <0.5	549	298	5.354	0.72333 ± 35	6.18	5.23	23	0.7216	22 ± 2
262 0.5-0.9	542	297	5.311	0.72402 ± 24	4.19	7.30	32	0.7216	30 ± 2
262L <0.5 1 M HCl	21	150	0.406	0.72132 ± 19	—	—	—	—	—
262R <0.5 1 M HCl	552	49	36.65	0.73284 ± 21	—	—	—	—	—
263 <0.5	408	321	3.707	0.73855 ± 54	4.06	10.54	24	0.7373	65 ± 4
263 0.5-0.9	418	229	5.316	0.74001 ± 27	4.64	14.03	36	0.7373	76 ± 4
263L <0.5 1 M HCl	13.7	162	0.265	0.73674 ± 74	—	—	—	—	—
263R <0.5 1 M HCl	485	66	21.32	0.74642 ± 42	—	—	—	—	—
264 0.5	264	162	4.752	0.72378 ± 63	4.13	22.26	97	0.7172	133 ± 10
264 0.5-0.9	235	175	3.916	0.72257 ± 16	4.02	27.17	97	0.7172	165 ± 9
264L <0.5 1 M HCl	132	41	9.292	0.73093 ± 96	—	—	—	—	—
264L 0.5-0.9 1 M HCl	132	21	15.74	0.73889 ± 72	—	—	—	—	—
264R <0.5 1 M HCl	186	196	2.759	0.72107 ± 80	—	—	—	—	—
264R 0.5-0.9 1 M HCl	173	244	2.061	0.72045 ± 65	—	—	—	—	—

Samples 3 and 4 sheared Upper Triassic limestone, $47^{\circ}33'23''$ N $13^{\circ}23'30''$ E; 5, undeformed Triassic limestone adjacent to the fault, $47^{\circ}33'17''$ N $13^{\circ}23'51''$ E; Kreuzeck fault gouge 263, $46^{\circ}49'36''$ N $13^{\circ}4'1''$ E; and fault gouge 262 and adjacent garnet mica schist 264, $46^{\circ}51'22''$ N; $13^{\circ}3'59''$ E (Drassnitzbach diversion tunnel 1,107 and 830 m, respectively). Sample suffix L Amb indicates carbonate cation-exchange leachate; L, 1 M and 3 M HCl leachate; R, residue.

* Maximum error 1.5%.

conditions which are suitable for forming and preserving minerals in faults. For example in the San Andreas fault system expandable clays, in contrast to normal sedimentary environments, are assumed to be preserved down to a depth of 10 km (see ref. 6). However, 2 M mica modifications and anchimetamorphic illite crystallinity values indicate that clay mineral formation occurred at very-low-grade metamorphic temperatures ($150\text{--}270^{\circ}\text{C}$).

Although in both examples the geological time brackets are wide and the Rb-Sr ages, with exception of the Pailwand fault, do not fulfil the requirements of an independent isochron, the combination of these techniques applied in this study is a promising tool to elucidate the timing of faulting in low-temperature shear zones which has been impossible by other means. The ages obtained from the cessation of tectonic faulting activities are consistent with the general geological information known from these areas.

Steep NW-SE striking fault zones oblique to the N-S direction of compression in the Pailwand area just postdate Upper Jurassic-earliest Cretaceous gliding and/or thrusting at anchimetamorphic conditions and clearly exclude Tertiary reactivation.

Uplift in the tectonic Tauern Window started in the Oligocene fairly soon after the metamorphic climax (>35 Myr)¹⁵. Contemporaneous movement along discrete fault zones must have taken place in the laterally situated Kreuzeck Group. Additional signs of this tectonic activity are granitic magmas (Wöllanerkopf granite 31 Myr)¹⁰ and mafic dykes ($24.2\text{--}40.6$ Myr)¹⁶ in this area having triggered simultaneous fluid circulation along major fault zones. These magmatic bodies are contemporaneous with intrusions along the important Periadriatic lineament 20 km to the south.

The results of this study point to a potential general application of these techniques. They may help to establish the timing of (1) new mineral formation in fractures around planned waste disposal sites, (2) the formation of mineral deposits in low temperature hydrothermal systems. The method might be helpful in tectonic earthquake research as well as in practical engineering geology projects.

The research was funded by grants from the Austrian National Bank and the Fonds zur Förderung der Wissenschaftlichen Forschung. We thank Drs Litscher and Jacobi for providing samples, Professor Frank for laboratory support, Ms Jelence for helping in isotope chemistry and Mr Leitner for drafting the figures. Former drafts of this paper benefited from reviews by S. Chaudhuri, W. Frank and L. Hoke.

Received 17 November 1986; accepted 9 March 1987.

- Kamineni, C. D. & Stone, D. *Contr. Miner. Petrol.* **83**, 237-246 (1983).
- Lyons, J. B. & Snellenburg, J. *Geol. Soc. Am. Bull.* **82**, 1749-1752 (1971).
- Downey, W. S. & Mitchell, J. G. *Geophys. J. R. astr. Soc.* **65**, 247 (1981).
- Bonhomme, M. G., Bühlmann, D. & Besnus, Y. *Geol. Rdsch.* **72**, 105-117 (1983).
- Riedmüller, G. *Tschermaks miner. petrogr. Mitt.* **25**, 219-242 (1978).
- Wu, F. T., Blatter, L. & Roberson, H. *Pure appl. Geophys.* **113**, 87-95 (1975).
- Wu, F. T. *Pure appl. Geophys.* **116**, 655-689 (1978).
- Tollmann, A. in *The Geodynamics of the Eastern Alps* (eds Flügel, H. & Faupel, P.) 112-125 (Deuticke, Vienna, 1987).
- Kralik, M., Krumm, H. & Schramm, J. M. in *The Geodynamics of the Eastern Alps* (eds Flügel, H. & Faupel, P.) 164-178 (Deuticke, Vienna, 1987).
- Frank, W., Kralik, M., Thöni, M. & Scharbert, S. in *The Geodynamics of the Eastern Alps* (eds Flügel, H. & Faupel, P.) 272-281 (Deuticke, Vienna, 1987).
- Kralik, M. *Tschermaks miner. petrogr. Mitt.* **32**, 49-67 (1983).
- Kralik, M. *Geochim. cosmochim. Acta* **48**, 527-533 (1984).
- Kübler, B. *Bull. Cent. Rech. Pau* **2**, 385-397 (1968).
- Clauer, N. *Chem. Geol.* **27**, 115-124 (1979).
- Cliff, R. & Droop, G. T. R. *J. metamorphic Geol.* **3**, 403-415 (1985).
- Deutsch, A. *Contr. Miner. Petrol.* **85**, 45-57 (1984).
- McIntyre, G. A., Brooks, C., Compston, W. & Turek, A. *J. geophys. Res.* **71**, 5459-5467 (1966).

Borehole evidence for a thick layer of basal ice in the central Ronne Ice Shelf

Hermann Engelhardt & Jürgen Determann

Institut für Geophysik, Westfälische Wilhelms-Universität,
Corrensstrasse 24, D-4400 Münster, FRG

The Ronne Ice Shelf, floating between Berkner Island and the Antarctic peninsula, is the biggest ice shelf in Antarctica. It drains an estimated $1.2 \times 10^6 \text{ km}^2$ of the Antarctic ice sheet, much of it resting on bedrock below sea level. Consequently, the balance and dynamics of this ice shelf is of importance to Antarctic glaciology, especially with regard to the integration of this part of the cryosphere into the global processes that control the climate of the Earth. Extensive radio-echo sounding (RES) by Robin and others revealed reflections in the central part of the Ronne Ice Shelf at the relatively shallow depth of 100–200 m below surface¹. The interpretation of these echoes, which varied in strength, was ambiguous, and the possibility of internal reflecting horizons was thoroughly discussed. But after surface elevation measurements by radar altimeter from drifting balloons appeared to fit the presence of thin ice, it was decided to base a thickness map of the Ronne Ice Shelf on these RES echoes^{1–4}. We now present direct observational evidence from boreholes that the total ice thickness is much greater than mapped, and that the shallow RES reflections therefore do come from internal horizons.

One objective of the German Antarctic Expedition in the summer of 1985/86 was to carry out a 900-km transverse on the Ronne Ice Shelf in order to study the nature of the central section of the ice shelf (Fig. 1). Two of the six-member party did the surveying, one was responsible for taking snow samples and temperature measurements down to 10 m at each measurement site, one was a mechanic who took care of the vehicles of our 40-t convoy, and the rest were in charge of the RES, seismic and drilling programme. Equipment included an upgraded hot-water drill of the type designed by Taylor⁵ and used successfully on Variegated Glacier⁶. It was used here in cold ice for the first time. An initial drilling speed of 120 m h^{-1} was achieved in the upper ice.

The drilling site (HC) at $78^\circ 18' 31'' \text{ S}$ and $57^\circ 07' 39'' \text{ W}$ was located in the apparently shallow central area of the ice shelf where our RES reflection came from a depth of only 170 m in good agreement with the depth at this point shown by the map of ice thickness^{1–3}. After two attempts to penetrate the ice shelf, the ice shelf bottom was reached at a depth of 465 m, on 31 January 1986 at 0:15 h, after 12 h of uninterrupted drilling. The depth was close to the 500 m total length of the drilling hose.

The breakthrough was indicated by a change in water level in the borehole, by an increase of hose tension, and by the drill lowering at higher speed. On lifting the drill from sub-basal water, a definite resistance was felt when the drill encountered the edge of the mouth of the borehole. This effect was reproduced several times and taken to indicate the bottom of the ice shelf. The borehole was reamed to a diameter of 10 cm down to a depth of 100 m in one hour. From this performance we estimate that the initial borehole had an average diameter of about 7 cm. Owing to lack of time no inclinometry could be done, but from our experience with hose tension, which was monitored continuously during drilling, we expect the borehole to be practically vertical.

When drilling below a depth of 380 m, the borehole bottom was tested every 5 m by lowering the drill faster than maximum drilling speed. This test indicated a hard definite bottom down to a depth of 430 m. Within the lowest 35 m of the borehole the bottom became gradually softer; the drill did not seem to stop

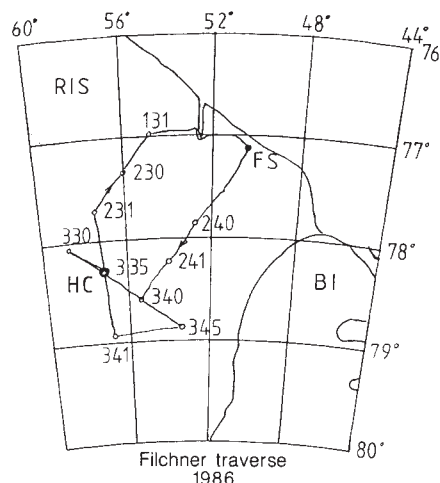


Fig. 1 Traverse route on the Ronne Ice Shelf (RIS), starting from Filchner Station (FS) along marked and surveyed measurement sites to the drilling camp, at site 335 west of Berkner Island (BI). Return via sites 231, 230 and 131. The numbered sites are part of an overall $50 \times 50 \text{ km}$ reference grid of the RIS programme.

abruptly at a solid water-ice boundary, but seemed to slide into slush that stopped the drill more and more smoothly as it approached the very bottom of the ice shelf. Also the drilling speed, which gradually decreased to 30 m h^{-1} at a depth of 430 m, because of cooling of the hot water in the increasing length of submerged hose, suddenly could be increased again, reaching 60 m h^{-1} at the bottom, because an ice-water mixture requires less heat to be penetrated by melting than solid ice.

The water level in the borehole was initially at a depth of 35 m, slowly rising to 30 m during drilling. When the borehole pierced the ice shelf, the water level dropped to 53.77 m below the surface. This water level corresponds to a 411-m column of nearly fresh water standing in the borehole and having a density of close to $1,000 \text{ kg m}^{-3}$. If we were to replace this fresh-water column with sea water of an average density of 1027.8 kg m^{-3} as found at the front of the ice shelf⁷, the water level would drop to 65 m corresponding to the height of the ice shelf surface above sea level.

Below the RES horizon at a depth of 170 m there is a layer of ice 295 m thick. Our drilling experience from boreholes through the cold Ekstöm Ice Shelf did not show an effect of ice getting soft at depth because of approaching the melting point; therefore, we conclude that in its lowest 35 m the ice is not fully consolidated and may consist of slush. A soft basal layer also would explain the extreme weakness of seismic reflections that we and the British Antarctic Survey party (B. Herrod, personal communication) obtained from the ice-shelf bottom in this area. Slush may also have floated up into the borehole, causing a temporary hang-up of the drill when it was pulled out.

Because the freezing point of water increases with decreasing pressure, ice will crystallize in water that is depressurized at its freezing point. Therefore, the origin of the lower 295 m of the ice shelf could be frazil ice, that is, needles or platelets of ice nucleated and grown in cold water that is being moved upward by ocean currents flowing along a sloping ice shelf bottom or sea bed and thus under decreasing pressure⁸. The ice shelf thins as it moves away from its grounding line. Our seismic data show the sea floor at a depth of 770 m below the snow surface in the vicinity of the borehole. Towards the front of the ice shelf the sea floor rises to a depth of less than 300 m⁷. Ocean currents flowing along these slopes under the central Ronne Ice Shelf have the potential to produce frazil ice. The ice accreting at the base of the ice shelf would be saline and at the freezing temperature when it formed. As the frazil ice is thickening underneath, it is made denser by compaction and recrystallization. Because of segregation in the freezing process, and because of

mechanisms that let the heavier brine sink in an ice-sea-water mixture, the salinity should be low and should increase downward. Our RES measurements near the front of the ice shelf, where both the internal reflection and the reflection from the real ice shelf bottom are visible, show that the basal echo fades rapidly as the basal ice layer gets thicker further south. This phenomenon, which was also detected on earlier records of the Cambridge RES group (C. Doake, personal communication), indicates, if one assumes a constant reflection coefficient at the melting and therefore smooth ice-water interface near the front, that absorption of RES energy in the lower part of the ice is high, which accords with the existence of slightly saline ice.

RES profiles south of the partially grounded area between Henry and Korff Ice Rises, now known as the Doake Ice Rumples (DIR), show strong bottom echoes from ice 1,000 m thick¹. Consequently, the internal, normally much weaker reflections are absent south of the DIR. We take this to indicate that there is no saline ice layer before the ice moves over the DIR and before it thins as it flows towards the drill site at HC. Thus, the saline ice could only be formed after the shelf has moved away from the partially grounded area of the DIR, which is located approximately 260 km upstream of the borehole site². The internal RES layer is probably caused by disturbances at the bottom as the ice moves over the DIR and through the hinge zone at the grounding line. If we assume that the bottom crevasses do not penetrate deep into the ice from below, because it is mechanically difficult and because it has not been observed in the Ross Ice Shelf in similar situations, then, at HC most of the 295 m of ice has accreted below the internal horizon. Use of an average speed of 400 m per year for the ice shelf from the grounded area to HC, estimated from available velocity measurements (H. Kock and A. Wigand, personal communication), gives an average ice accretion rate of 45 cm per year at the base of the ice shelf. Strong, but not unrealistic ocean currents are needed to account for these accretion rates. Nothing is known about them under the ice shelf, but at the ice-shelf front strong ocean currents of up to 0.4 m s^{-1} were observed⁷. For example, a column of water only 4 m deep, moving at this speed along a slope of 1 m km^{-1} would produce the inferred accretion rate for frazil ice. This bottom accretion rate is much greater than the surface accumulation rate of around 10 cm yr^{-1} ⁹. The mass balance of the central Ronne Ice Shelf is controlled mainly by freezing processes underneath the ice shelf. The warm basal temperature together with the greater accretion rate of the basal ice will make the mean temperature of the ice shelf considerably warmer than that of an ice shelf dominated by surface accumulation. Such a temperature profile has been measured on the Amery Ice Shelf¹⁰, where basal accretion of layer 160 m thick of saline ice has been determined by Morgan¹¹. Accretion rates of around 70 or 30 cm per year have been deduced from the data^{10,11}. Because of the existence and nature of the newly discovered lower ice with a volume of approximately $15,000 \text{ km}^3$, the spreading rate is enhanced and the overall dynamics of the Ronne Ice Shelf is greatly altered. Our findings point to the need for further studies, such as temperature profiles within the ice and ocean currents beneath the ice shelf. These are necessary for the full understanding of the regime of the Ronne Ice Shelf.

We acknowledge Gordon Robin and Christopher Doake and their co-workers whose work greatly influenced our decision to drill at HC and thank C. Bentley, C. Doake and G. Robin for discussions, and our four traverse colleagues N. Beck, H. Kock, G. Miller and A. Wiegand for support. H. Ruffi solved many engineering problems and M. Hoyer helped in preparing Antarctic equipment and testing the hot-water drill. RES equipment was developed by O. Schomburg under Professor F. Thyssen. We are grateful to the Alfred-Wegener-Institut für Polarforschung for logistical support and the Bundesministerium für Forschung und Technologie and the Deutsche Forschungsgemeinschaft for financial support.

Received 3 November 1986; accepted 9 March 1987.

1. Robin, G. de Q. *et al.* *Nature* **302**, 582-586 (1983).
2. Crabtree, R. D. & Doake, C. S. M. *Ann. Glaciol.* **8**, 37-41 (1986).
3. Drewry, D. J. (ed.), *Glaciological and Geophysical Folio Sheet 4* (Scott Polar Research Institute, Cambridge, 1983).
4. Levanon, N., Julian, P. R. & Suomi, V. E. *Nature* **268**, 514-516 (1977).
5. Taylor, P. L. *Ice Drilling Technology, U.S. CRREL Special Report* (U.S. Army Cold Region Research and Engineering Laboratory, Hanover, New Hampshire) 84-34; 105-117 (1984).
6. Kamb, B. *et al.* *Science* **227**, 469-479 (1985).
7. Gammelsrod, T. & Slotsvik, N. *Polarforschung* **51**, 101-111 (1981).
8. Doake, C. S. M. *Polar Rec.* **18**, 37-41 (1976).
9. Reinwarth, O. & Graf, W. *Filchner-Ronne Ice Shelf Programme Report No. 2* (compiled by H. Kohnen) 7-17 (Alfred Wegener Institute for Polar Research, Bremerhaven 1985).
10. Robin, G. de Q. *The Climatic Record in Polar Ice Sheets*, 97; 121-122 (Cambridge University Press, 1983).
11. Morgan, V. I. *Nature* **238**, 393-394 (1972).

Holes in a stearic acid monolayer observed by dark-field electron microscopy

N. Uyeda, T. Takenaka, K. Aoyama, M. Matsumoto & Y. Fujiyoshi

Institute for Chemical Research, Kyoto University, Uji, Kyoto 611, Japan

The structure and behaviour of surface monolayers have been extensively studied for the past few decades, leading to practical applications to electronic devices as well as to biomimetic study. The basic concept underlying these wide applications is thought to be the homogeneity expected of monomolecular films properly compressed on the water surface. The structures of surface films have been investigated by spectroscopy, X-ray and electron diffraction¹⁻³ and other indirect methods. The real structure, however, should have been inspected using direct electron microscopy. Here we report the first practical observation which clearly reveals the surprising feature that the monomolecular film of stearic acid is not completely homogeneous but includes densely scattered small holes, which is contrary to the currently accepted concept of homogeneity.

Ries *et al.*⁴⁻⁶ have been studying the mechanism of film formation and collapse by electron microscopy, using specimens which have been shadow-casted to enhance the contrast. Although discrete features such as island formation were revealed during the compression, no fine structures were detected. Another set of observations without shadows has been performed by Baumeister *et al.*⁷ and Fryer *et al.*⁸. However, no information on the general features of thin films has been given since the bright-field mode was adopted in view of the high-resolution observation based on the phase contrast mechanism. Here, we adopted the dark-field imaging mode which was found to give satisfactory image contrast for surveying the appearance of films at rather low magnification. The minimum dose system was also used to reduce the radiation damage on organic molecules such as stearic acid.

A solution of stearic acid ($1.66 \times 10^{-3} \text{ mol kg}^{-1}$) was prepared by dissolving the solute of four-nine grade in hexane of spectroscopic grade. The monomolecular film was spread on the water surface at 20°C in a Teflon-coated trough. The film compression was started after 30 min so that the solvent molecules were completely evaporated. The compression speed was $8.0 \text{ cm}^2 \text{ min}^{-1}$, and the surface pressure was recorded by the Wilhelmy method. The pressure(π)-area(A) isotherm begins to rise at $0.22\text{--}0.23 \text{ nm}^2$ per molecule and there follows a rapid increase in pressure (Fig. 1). The film has a transition point near a pressure of 20 mN m^{-1} and collapses at about 50 mN m^{-1} . This behaviour is similar to past observations reported by Ries⁴ and also by Adam⁹.

The specimens for electron microscopy were prepared by a

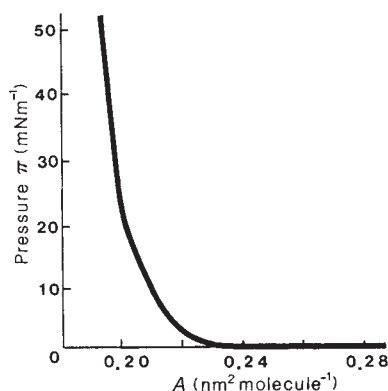


Fig. 1 Pressure (π)-area (A) isotherm for stearic acid monolayer at 20 °C.

horizontal attachment technique at approximately ten sequential stages during the compression. The small patches of film were transferred to specimen grids covered with thin carbon supporting films so that the hydrophobic groups faced inwards. The transfer ratio measured at 15 mN m⁻¹ was found to be 100% (within experimental error). The adhesion of the hydrocarbon chain to the hydrophobic carbon film was found to be so good that the electron microscope features in general did not show any conspicuous differences even after the specimen was washed with water. The practical observation was based on the dark-field imaging performed by shifting the objective aperture to block the main electron beam and at the same time to admit the electrons scattered from the regular or irregular arrays of chain molecules. Without the main beam, the monolayer appears as bright images on a dark background. As an aliphatic compound such as stearic acid is known to be very radiation sensitive, the use of a minimum-dose system was indispensable to reduce the damage by electron irradiation. For further protection, the overall observation was performed at a magnification not higher than 4,000. The selected-area electron diffraction patterns were also recorded to determine the molecular arrangements in the surface films.

The series of electron photomicrographs thus obtained show that irregular aggregates of two-dimensional fine particles of about 50 nm in diameter had been formed in a misty background even before the compression was started. The configuration of monolayer changed with the compression without showing any rise in the surface pressure. Figure 2a shows a typical feature of the film which appeared when the area $A = 0.36$ nm² per molecule. The dendritic chains of aggregates in different sizes form a net-like structure holding a small vacant area among them. When the surface pressure became measurable, the net structure turned into large islands surrounded by wide openings which gradually decreased as the pressure increased, finally leaving a coherent film when the surface pressure exceeded 0.5 mN m⁻¹. The electron micrographs of the film, compressed to 30 mN m⁻¹ are shown in Fig. 2b, where the film assumes a remarkable two-dimensional coherency. Unexpectedly, however, the film contains small holes uniformly scattered all over the surface area. The holes are not round but assume a somewhat irregular shape and their sizes range from 300 to 50 nm. The mean diameter is ~ 100 nm, whereas the average density of the holes was estimated to be about 40–50 per μm^2 . These holes have already appeared in the large-size islands formed at zero surface pressure. Once the holes have been formed the general appearance is thought to hold without any apparent variation in hole distribution until the continuous film is formed and compressed until it collapses. The relationship to the transition point was not found in these observations.

The serial observation of electron diffraction indicated that some ordering takes place after the film has attained coherence

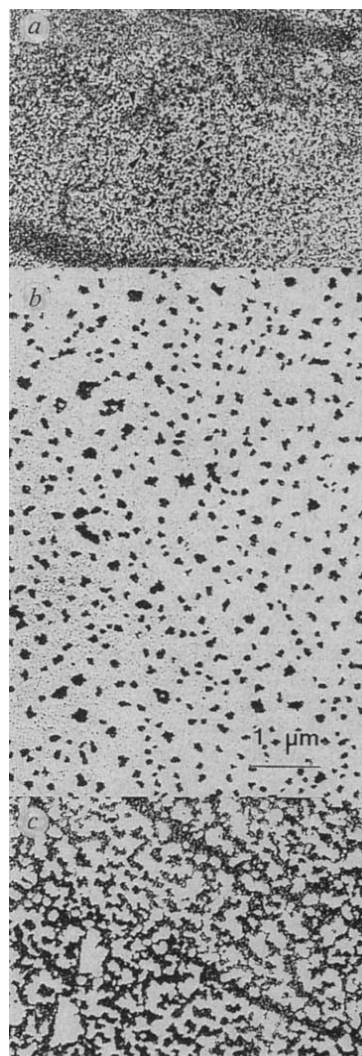


Fig. 2 Dark-field electron micrographs of stearic acid monolayer. *a*, Dendritic net of aggregates at early stage of compression ($\pi = 0$ mN m⁻¹, $A = 0.36$ nm² per molecule). *b*, Coherent monomolecular film containing small irregular holes ($\pi = 30$ mN m⁻¹, $A = 0.18$ nm² per molecule). *c*, Flaky configuration of film collapsed over $\pi = 52$ mN m⁻¹.

regardless of the presence of the small holes although the lattice orientation is present in the earlier stage of island formation. A typical diffraction pattern assumes a hexagonal symmetry as shown in Fig. 3a for an example obtained from a continuous part of the monolayer. Because the diffraction spots are arc-shaped and higher-order reflections are not observed clearly enough, the crystallinity is thought to be rather poor. By the analysis of elliptical diffraction patterns obtained by tilting the specimen for deeper glancing angles, it was confirmed that the film is clearly a monolayer with the molecules closely packed in a two-dimensional hexagonal or pseudo-hexagonal lattice whose unit-cell length is 0.481 nm. When one molecule is assigned to the unit cell, the surface area occupied is estimated to be 0.20 nm² per molecule which is slightly smaller than that found from the extrapolation of the pressure isotherm. When the surface pressure exceeds 45–50 mN m⁻¹, the monolayer became a discrete film as shown in Fig. 2c, where the small irregular patches of flakes have gained much contrast. The electron diffraction indicated that the film was not monomolecular any more but had become a three-dimensional configuration again with hexagonal or pseudohexagonal symmetry.

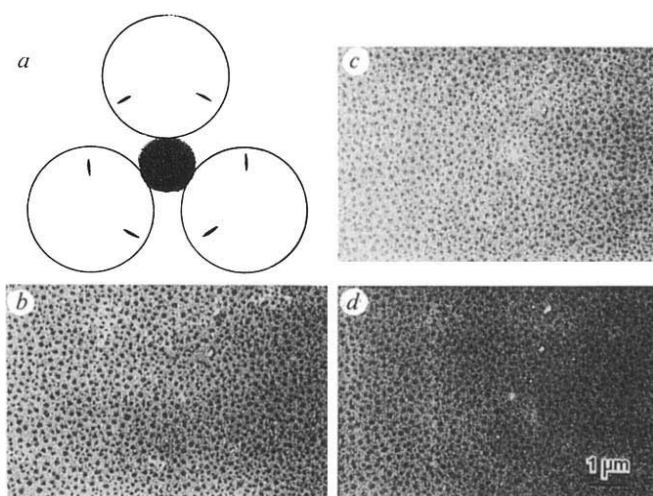


Fig. 3 Electron diffraction from monolayer and dark-field image formation. *a*, Diffraction pattern for normal beam incidence. The circles indicate the position of objective aperture for dark-field imaging. *b*–*d*, Dark-field images of the same area, each formed by different pair of major reflections.

Although the dark-field observation of the large island or coherent film was performed mostly by allowing a pair of neighbouring diffraction spots through the objective aperture as circled in Fig. 3*a*, it is worthwhile confirming that these images actually reveal the overall appearance of the target area. Such an extremely thin crystal as the monolayer of a two-dimensional or nearly two-dimensional lattice is known to cause wide-angle simultaneous reflections so that the electron waves scattered by the individual atoms or molecules evenly propagate into all sets of equivalent reflections. This means that whichever reflection is used, the same specimen area is imaged regardless of the presence of the misaligned area consisting of, for example, arrays of differently tilted molecules due to a slight bending of the film. An observation to confirm these results was performed in the present case by taking three dark-field images of the same area using each of three pairs of reflections as circled in Fig. 3*a*. The three observations resulted in the formation of three identical images as shown in Fig. 3*b*, *c* and *d*. The above result indicates that imaging with only a single pair of reflections or even a single reflection is sufficient to survey the overall features of the target area without any localization.

The above finding does not necessarily exclude the possibility that the dark areas are not completely vacant but contain certain amount of chain molecules which assume some orientation different from that of the matrix. If present, however, these molecules, especially the constituent carbon atoms, are expected to scatter strong incoherent electron waves which pass through the aperture to give a certain brightness to the corresponding areas. In view of their darkness, the actual number of molecules, if any, is considered to be rather small, leaving those areas almost vacant.

As for the presence of small holes, the possibility of molecular evaporation might be considered as the specimen has to be placed in the high vacuum of the microscope column. To examine this possibility, two photographs were recorded for the same image field, one immediately after the specimen was set in the column and the other after leaving it in the vacuum intact for an additional 90 min. A close examination of the two showed that no differences were detected especially in the shape and size of the individual holes as well as in their distribution. This result indicates that the molecular adhesion to the carbon film is considerably strong so that even the high vacuum does not cause any detectable variation.

Holes seem already to exist in much earlier stages of

monolayer formation as seen in Fig. 2*a*. Some of those small thin areas indicated by arrows may remain to form more obvious holes when the dendritic networks grow into large discrete islands. The rise in the surface pressure may be attributed mainly to the compression of these islands together into the final coherent monolayer still including the original irregular holes.

Note added in proof: Although invisible, two dimensional networks of flatly lying chain molecules may cover the hole. As seen in Fig. 2*b*, some larger holes have a small inner bright spot which is considered to be a tiny fragment of regular monolayer surrounded by these networks.

Received 1 October 1986; accepted 27 February 1987.

1. Havigna, E. & de Vael, J. *Recl. Trav. Chim.* **56**, 375–381 (1937).
2. Germer, L. H. & Storks, K. H. *J. chem. Phys.* **6**, 280–294 (1938).
3. de Vael, J. & Havigna, E. *Recl. Trav. Chim.* **59**, 770–776 (1941).
4. Ries, H. E. Jr *Scient. Am.* **211** (1961).
5. Ries, H. E. Jr, Matsumoto, M., Uyeda, N. & Suito, E. *Adv. Chem. Ser.* **144**, 286–293 (1975).
6. Ries, H. E. Jr, Matsumoto, M., Uyeda, N. & Suito, E. *J. Colloid Interface Sci.* **57**, 396–398 (1976).
7. Baumeister, W. & Hahn, M. H. *Nature* **241**, 445–446 (1973).
8. Fryer, J. R., Hann, R. A. & Eyres, B. L. *Nature* **313**, 382–384 (1985).
9. Adam, N. K. *The Physics & Chemistry of Surfaces* 43–51 (Dover, New York, 1968).

Assembly of micelles into higher-order structures with increasing salt concentration

J. Tabony*, A. de Geyer* & L. F. Braganza†

* Département de Physico-Chimie, CEN Saclay, 91191 Gif-sur-Yvette, France

† Institut Laue-Langevin 156X, 38042 Grenoble Cedex, France

Concentrated aqueous surfactant solutions of ~50% by weight frequently form liquid crystalline mesophases^{1,2} having repeat distances of ~50 Å. At lower surfactant concentrations, however, the ordering is lost and micelles³ are formed. We present neutron scattering and nuclear magnetic resonance observations of dilute (~1–10% by weight) aqueous ionic surfactant/co-surfactant solutions which show that upon the addition of salt, disperse micellar aggregates assemble into mesophases of higher order which have repeat distances of up to 1,000 Å. This raises the question as to the structure of these phases and the origin of the forces responsible for ordering over such large distances. Studies of this assembly type of behaviour may provide insight into the manner in which mixtures of simple surfactant-type molecules might form interconnected network structures such as are found in biological systems.

For this study we used mixtures of 1–10% by weight each of surfactant and co-surfactant dissolved in salt solutions. The phase behaviour of such systems has recently been described by Benton and Miller⁴, who found that progressive addition of salt to an aqueous micellar dispersion (I) results in the formation of firstly an optically birefringent (L) phase followed by another optically isotropic (C) mesophase at higher salt concentrations. Hence the interconversion between different structures having the same overall composition can be achieved simply by changing the ionic strength. The surfactant and co-surfactant we employed were tetradecyl trimethyl-ammonium bromide and pentanol respectively. For ~2% each by weight of surfactant and co-surfactant I, L and C mesophases were formed at salinities of ~0%, 2% and 3% by weight respectively of NaBr in D₂O. At these concentrations, all three phases had viscosities similar to water. Salinity ranges in which the different mesophases occurred were critically dependent on the concentrations and relative proportions of surfactant and co-surfactant.

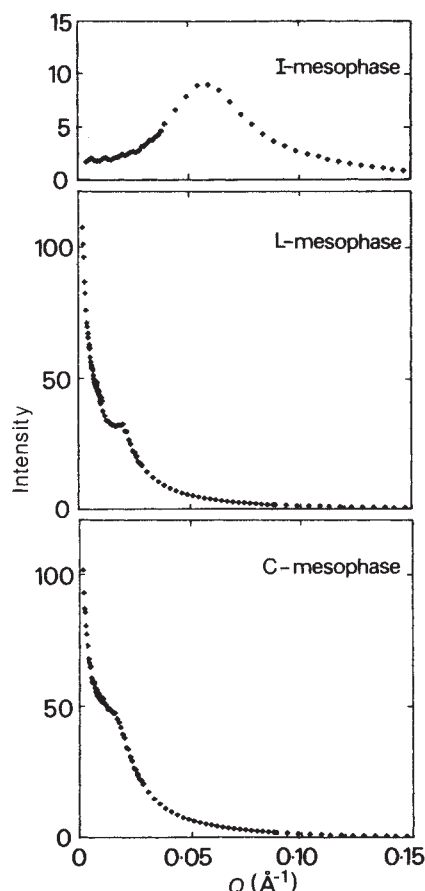


Fig. 1 Neutron small-angle scattering patterns of the I, L and C mesophases obtained by changing the ionic strength. The samples were composed of 0.54 g tetradecyl trimethylammonium bromide and 0.50 g pentanol in 10 ml D_2O NaBr solutions. For these amounts of surfactant and co-surfactant the I, L, C mesophases were obtained at salt concentrations of respectively 0%, 2% and 6% by weight in D_2O . The peak positions correspond to a spacing of ~ 300 Å for the L mesophase and ~ 400 Å for the C mesophase. The spectra were recorded on the D11 and D16 diffractometers at the Institut Laue-Langevin, Grenoble, France. The intensity scale is in arbitrary units and $Q = (4\pi/\lambda) \sin \theta$ where λ is the neutron wavelength and 2θ the scattering angle.

Figure 1 shows neutron small-angle scattering patterns for the I, L and C mesophases. All measurements were done at $(23 \pm 1)^\circ C$. The I phase shows a broad maximum typically obtained for aqueous charged micellar solutions⁵. Unlike the I phase, the L and C mesophases give peaks at much lower values of reciprocal space, which for the cases illustrated correspond to spacings of ~ 300 Å and ~ 400 Å for the L and C mesophases respectively. A common feature of all three spectra is that the scattering extends to reciprocal space values of ~ 0.1 Å⁻¹; this is attributed to scattering from bilayers containing surfactant and co-surfactant, approximately 30 Å thick, which constitute the common building block for the different mesophases.

Figure 2 shows deuterium NMR spectra of samples in which the surfactant trimethylammonium polar head group had been selectively deuterated. In contrast to the single resonance observed for the isotropic C phase, a doublet arises for the L mesophase due to the spontaneous alignment in the magnetic field of bilayers containing the surfactant molecules⁶. Figure 3 shows the distribution of neutron counts over a two-dimensional detector for an L mesophase oriented in a similar way. Intensity maxima are observed perpendicular to the applied field (H) direction which, for the sample illustrated, correspond to a mean

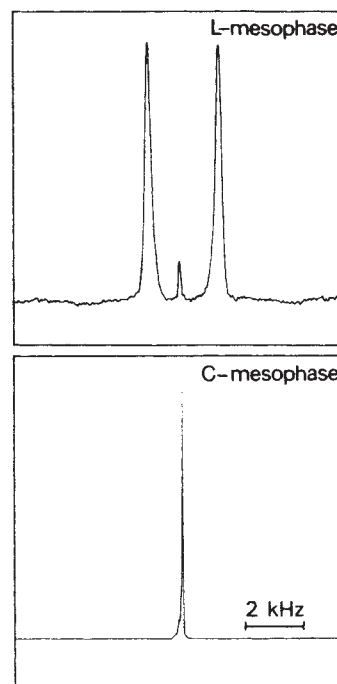


Fig. 2 Deuterium NMR spectra of the L and C mesophases prepared with surfactant molecules deuterated at the trimethylammonium polar headgroup. Samples were composed of 1.24 g $C_{14}H_{29}N^+(CD_3)_3Br^-$ and 1.18 g pentanol in 10 ml of 1% and 2% by weight NaBr H_2O solutions for the L and C phases respectively. The doublet observed for the L mesophase is due to alignment in the magnetic field of structures containing the surfactant molecules. The weak central peak is attributed to surfactant molecules below the critical micelle concentration. The field strength was 11.7 T and the operating frequency 76.7 MHz.

spacing of ~ 510 Å. The strong magnetic orientation can only arise if the objects are highly anisotropic in shape. Thus other possibilities such as the occurrence of spheroidal vesicles can be excluded. Diffraction measurements on samples containing progressively higher concentrations of surfactant and co-surfactant resulted in large decreases of the spacing and the appearance of further reflections. At concentrations of $\sim 12\%$ by weight each of surfactant and co-surfactant, four equally spaced lamellar reflections are observed and the repeat spacing is ~ 90 Å. This diffraction pattern shows that the surfactant and co-surfactant molecules assemble into flat bilayer sheets or platelets. Furthermore, a Fourier synthesis with the diffraction amplitudes gave a picture of the neutron coherent scattering length density across the planar structures⁷ which showed the mixed surfactant/co-surfactant bilayers to be ~ 30 Å thick. The most dilute L mesophase possible to prepare gave a spacing of $\sim 1,000$ Å. This, and a more detailed analysis of the scattering pattern, suggests that the bilayer sheets or platelets have a length of $\sim 1,000$ Å. Provided the intersheet separation is less than the length of the objects, the resulting solution will not be isotropic but will behave as a lamellar nematic phase.

For the C phase, the spacings are larger by 20–25% than for the L phase at the same concentrations of surfactant and co-surfactant. Increasing the concentrations of surfactant and co-surfactant produced large decreases in the spacing but without the appearance of further reflections as for the L phase. The spacing decreased progressively from ~ 530 Å to 120 Å with increasing concentration of surfactant and co-surfactant in the range $\sim 3\%$ to 12% each by weight. Thus the structure expands and contracts with varying brine volume fraction. The observed spectra are not consistent with simulated scattering curves for

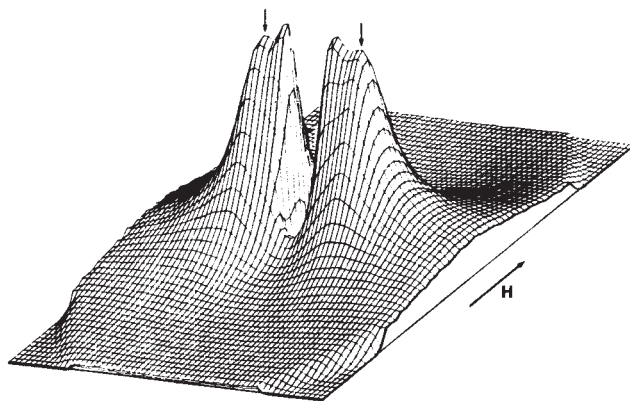


Fig. 3 The anisotropic distribution of neutron counts on the two dimensional 64×64 cm detector of the D17 diffractometer for a dilute L mesophase oriented in a magnetic field. The sample composition was 0.298 g tetradecyl trimethylammonium bromide, 0.270 g pentanol and 10 ml of 1.5% by weight NaBr in D_2O . Because of the slight solubility of pentanol in D_2O (~ 0.14 ml per 10 ml), the preparation of dilute L mesophases required a greater overall proportion of alcohol to surfactant than more concentrated samples. The intensity maxima (arrowed) on either side of the main beam correspond to a spacing of ~ 510 Å. Similar orientation was obtained for the most dilute samples with spacings of $\sim 1,000$ Å. The neutron wavelength was 16.0 Å and a sample-to-detector distance of 3.5 m was used. A beamstop prevents the main beam from hitting the central part of the detector. The sample was placed in a vertical magnetic field of ~ 6.5 T for a few minutes and transferred into the neutron beam for the measurement which lasted ~ 40 min. No loss of orientation was observed over this timescale. The direction of the magnetic field, H , with respect to the detector is shown by the arrow.

homogeneous or inhomogeneous particles with spherical symmetry⁸ such as a solution of large unilamellar vesicles⁹. Furthermore, the distinctive maxima seen in the simulated scattering curves for such objects do not show large positional changes on increasing the concentration of scatterers⁸. In addition, measurements whereby the internal scattering length density was modified by selectively deuterating the surfactant or co-surfactant did not affect the peak. This result shows that the peak does not arise from the internal structure of objects such as vesicles, since changes in their scattering length densities would give large changes in the peak intensity and/or position.

For curved surfaces, a plot of $Q^4 I(Q)$ versus Q (where $Q = (4\pi/\lambda) \sin \theta$, λ is the wavelength, 2θ the scattering angle and $I(Q)$ the scattered intensity) shows a maximum whose position in Q is related to the curvature¹⁰. The micellar I phase shows a maximum in the $Q^4 I(Q)$ versus Q plot which yields a curvature corresponding to a micellar radius of 20 Å. Consistent with a flat lamellar surface, the L phase shows no maximum. The C phase once again shows a maximum corresponding to regions of surface curvature of 20 Å. Such a curvature is not consistent with a large unilamellar vesicle-type structure.

The effect of adding salt seems to be twofold. First, it screens out the repulsive electrostatic interactions between small micellar aggregates thus encouraging them to come into contact and possibly fuse together so as to form larger aggregates of higher order. Second, it reduces the surfactant polar head group repulsions within the bilayer. The polar heads then approach one another and this results in a decrease of the local surface curvature¹¹. Progressive addition of salt to surfactant micelles in water will hence favour the formation of a dispersion of bilayer sheets where the curvature is zero. This explains the transition from the isotropic micellar I phase to the birefringent lamellar nematic L phase. Further addition of salt will, if possible, result in a local curvature of opposite sense to micelles. Thus, the addition of salt to a microemulsion can reverse the structure from oil-in-water droplets to water-in-oil droplets¹².

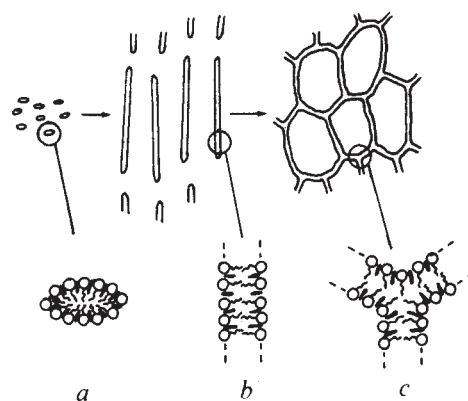


Fig. 4 Schematic sectional representations of *a*, the micellar I phase; *b*, the lamellar L phase and *c*, the C phase showing one way in which bilayers can fuse to form a network where the junctions have a surface curvature opposite to that of the micelles (Fig. 4). Moreover, consistent with observation, the curvature at the junctions would have a value similar to that in the micellar I phase. Such a structure would be expected to shrink and expand with changing brine volume fraction. It would also explain the proximity to the lamellar L phase and account for the increased spacing in going from the L to C phases. If the junctions were distributed in an orderly manner, then some long-range order would result and the interfacial layer might possess a surface topology similar to an infinite minimal periodic surface^{13,14}. Such surface topologies, which are thought to occur in some cubic liquid crystals and microemulsions^{15,16}, minimize the total surface free energy and might mean that these structures would be preferred to more random arrangements.

In the present case, no oil is present. However, one manner in which the net curvature can be partially reversed is by the formation of a three-dimensional interconnected network, resulting from the fusing of finite bilayer sheets or platelets, where the junctions have curvatures opposite to that of the micelles (Fig. 4). Moreover, consistent with observation, the curvature at the junctions would have a value similar to that in the micellar I phase. Such a structure would be expected to shrink and expand with changing brine volume fraction. It would also explain the proximity to the lamellar L phase and account for the increased spacing in going from the L to C phases. If the junctions were distributed in an orderly manner, then some long-range order would result and the interfacial layer might possess a surface topology similar to an infinite minimal periodic surface^{13,14}. Such surface topologies, which are thought to occur in some cubic liquid crystals and microemulsions^{15,16}, minimize the total surface free energy and might mean that these structures would be preferred to more random arrangements.

The above observations and arguments suggest that the long-range ordering observed here results from a balance of forces different from those in normal lyotropic liquid crystals and arise through a complicated and delicate interplay between the effects of local interfacial curvature combined with the reduction in electrostatic repulsions brought about by the gradual addition of salt.

Phase behaviour of this type is not uncommon for dilute aqueous solutions of surfactant and co-surfactant in the presence of salt⁴ and may have its counterpart in many biological systems. Eukaryotic cells, for example, have extensive three-dimensional intracellular membrane networks which divide the cytoplasm into specialized reaction vessels while also providing a large surface area for membrane-based biochemical processes¹⁷. The major constituent of biological membranes are lipids which are essentially surfactant-type molecules. Lipids mixed with other co-surfactant-type molecules in a suitable chemical environment could thus spontaneously form the initial framework of elaborate intracellular networks.

We thank our colleagues for helpful discussions and especially J. Torbet for assistance with the magnet.

Received 16 October 1986; accepted 6 March 1987.

1. Tiddy, G. J. T. *Phys. Rep.* **57**, 1-46 (1980).
2. Elkwall, P. in *Advances in Liquid Crystals* Vol 1 (ed. Brown, G. H.) Ch. 1 (Academic, New York, 1971).
3. Wennerstrom, H. & Lindmann, B. *Phys. Rep.* **52**, 1-86 (1979).
4. Benton, W. J. & Miller, C. A. *J. phys. Chem.* **87**, 4981 (1983).

5. Hayter, J. B. & Penfold, J. *JCS Faraday Trans. I* **77**, 1851 (1981).
6. Mantsch, H. H., Saito, H. & Smith, I. C. P. *Prog. NMR Spec.* Vol. 11 part 4, 211-272 (Pergamon, Oxford, 1977).
7. Worcester, D. L. in *Biological Membranes*, Vol. 3 (eds Chapman, D. & Wallach, D. F. H.) Ch. 1 (Academic, New York, 1976).
8. Glatter, O. in *Small Angle X-ray Scattering* (eds Glatter, O. & Kratky, O.) Ch. 5 (Academic, New York, 1982).
9. Lagner, P. in *Small Angle X-ray Scattering* (eds Glatter, O. & Kratky, O.) Ch. 10 (Academic, New York, 1982).
10. Kirste, R. & Porod, G. *Kolloid Zh.* **184**, 1 (1962).
11. Winsor, P. A. in *Liquid Crystals and Plastic Crystals* Vol. 2 (eds Gray, G. W. & Winsor, P. A.) Ch. 4 (Wiley, New York, 1974).
12. Winsor, P. A. *Trans. Faraday Soc.* **44**, 376 (1948).
13. Schoen, A. H. *Nasa tech. Note* (NASA, Washington DC, 1970).
14. Mackay, A. L. *Nature* **314**, 604 (1985).
15. Longley, W. & McIntosh, T. J. *Nature* **303**, 612 (1984).
16. Scriven, L. E. *Nature* **263**, 123 (1976).
17. Alberts, B. et al. in *Molecular Biology of the Cell*, Ch. 7 (Garland, New York, 1983).

Evidence for neuromelanin involvement in MPTP-induced neurotoxicity

Robert J. D'Amato*†‡, Guillermo M. Alexander§,
Robert J. Schwartzman§, Cheryl A. Kitt||,
Donald L. Price*|| & Solomon H. Snyder*†‡¶

* Departments of Neuroscience, † Pharmacology and Molecular Sciences, ‡ Psychiatry and Behavioral Sciences and || Department of Pathology, The Johns Hopkins University Medical Institutions, 725 North Wolfe Street, Baltimore, Maryland 21205, USA
§ Department of Neurology, Jefferson Medical College, Philadelphia, Pennsylvania 19107, USA

Exposure to 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP) reproduces certain clinical, pathological, and neurochemical features of Parkinson's disease¹⁻⁷. MPTP is metabolized by monoamine oxidase Type B to 1-methyl-4-phenylpyridine (MPP⁺)⁸, which is selectively accumulated by high-affinity uptake mechanisms into dopaminergic neurons⁹. Lyden *et al.*¹⁰ described low-affinity binding of MPTP to synthetic and retinal melanin. We showed that MPP⁺ binds to neuromelanin with high affinity^{11,12}, suggesting that in MPTP neurotoxicity, MPP⁺ enters nigral neurons by the dopamine uptake system and binds to neuromelanin, which serves as a depot, continuously releasing MPP⁺ until it destroys the cells¹¹. This model predicts that agents which compete with MPP⁺ binding to neuromelanin should partially protect the dopamine neurons from MPTP-induced toxicity. The most potent identified competitor for MPP⁺ binding to melanin is the antimalarial drug chloroquine¹², which has a high affinity for melanins¹³. In the present study, chloroquine, administered to monkeys in conventional anti-malarial doses before MPTP, protects them from MPTP-induced parkinsonian motor abnormalities, dopamine depletion in the striatum, and neuropathological changes in the substantia nigra.

Three monkeys received MPTP alone (0.35 mg per kg intravenously for 4 days); three monkeys received a 12-day pretreatment (short term) with 4 mg per kg chloroquine intramuscularly (i.m.) continued throughout the MPTP regimen and maintained for 10 days after the last injection of MPTP; three monkeys were pretreated with chloroquine for 24 days (long term).

The behaviour of the monkeys receiving MPTP alone was similar to that of four monkeys previously exposed to MPTP in our lab, and resembled that described in other reports¹⁴. On the fifth day after MPTP exposure the animals manifested decreased mobility and spontaneous movement, abnormal posture, rigidity of the neck and limbs, increased muscle tone, hyperactive reflexes, tremor of upper extremities, and lack of vocalization (Table 1).

Five of the six monkeys pretreated with chloroquine were

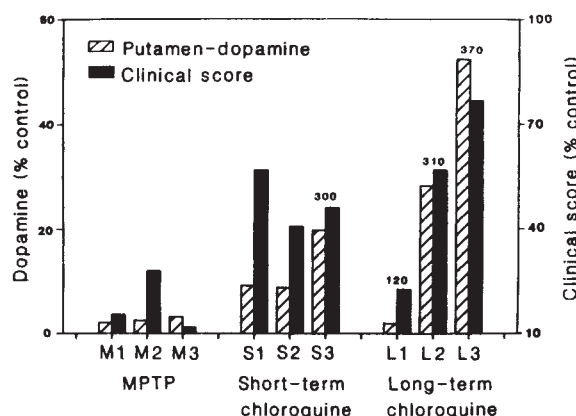


Fig. 1 Clinical and neurochemical status of monkeys treated with MPTP alone or pretreated with chloroquine for 12 days (S1-S3) or 24 days (L1-L3) prior to administration of MPTP. Clinical score and dopamine levels in putamen are expressed as percentage of control. Chloroquine levels (in ng ml⁻¹) are indicated above bars for monkeys S3, L1, L2 and L3. For details of measurements, see legends to Tables 1 and 2.

partially protected from MPTP-induced parkinsonian clinical symptoms. Of the three monkeys receiving long-term treatment with chloroquine, one animal (L-3) was almost completely protected except for a slight decrease in spontaneous movement. A second animal in this group (L-2) was also protected from the severe effects of MPTP. Although it exhibited modest rigidity, the monkey manifested no tremor, moved about the cage freely and vocalized extensively. However, one animal (L-1) demonstrated motor deficits as severe as monkeys receiving MPTP alone: the reasons for the failure of chloroquine protection are described below. All three animals receiving short-term chloroquine pretreatment were partially protected from MPTP neurotoxicity: although they had some rigidity, all moved about the cage readily, ate well, vocalized extensively and showed only modest tremor.

Results of neurochemical analyses closely paralleled the clinical findings (Table 2; Fig. 1). In monkeys receiving MPTP alone, amounts of dopamine, homovanillic acid (HVA) and tyrosine hydroxylase activity (TH) were markedly reduced in both the caudate and putamen. Dopamine was depleted to about 1% of control whereas HVA was at about 10% of control. The resulting increased HVA/dopamine ratio in the MPTP animals presumably reflects the greater turnover of dopamine in residual dopamine neurons. In monkeys receiving MPTP alone, TH immunocytochemical preparations⁷ revealed a pronounced reduction in the density of TH immunoreactive fibres and terminals in the putamen and to a lesser extent in the caudate nucleus as compared to controls (data not shown). The five chloroquine pretreated monkeys, which were clinically protected from MPTP neurotoxicity, showed much slighter reductions in levels of dopamine and TH, as well as TH-immunoreactive fibres and terminals, than the MPTP treated animals.

Neuropathological findings fitted well with the clinical and neurochemical observations. Representative neuromelanin-stained sections through the substantia nigra from each animal were rank ordered for cell loss by two naive observers whose rankings were identical and closely paralleled the neurochemical and clinical results. The correlation coefficient of the ranking with the caudate dopamine values (*R*) is 0.90. The greatest reduction in nigral cell number occurred in animals given MPTP alone (data not shown). Chloroquine pretreated animals had more surviving neuromelanin containing neurons, with the greatest number of cells remaining in the long term pretreatment group.

Thus, short-term treatment with chloroquine provided partial protection against clinical, neurochemical and neuropathologic

¶ To whom correspondence should be addressed.

Table 1 Neurological effects of chloroquine on MPTP neurotoxicity in monkeys

	Neurological status				
	Spontaneous movement	Tremor	Tone	Reflexes	Total (X5.26)
Control	10	5	2	2	100
MPTP					
M1	1	1	0.5	0.5	16
M2	1	1	2	1.5	29
M3	1	0	0.5	0	8
Short-term chloroquine/MPTP and MPTP					
S1	3	4	2	2	58
S2	3	3	1	1	42
S3	3	3	2	1	47
Long-term chloroquine/MPTP and MPTP					
L1*	0	2	1.5	1	24
L2	5	5	1	1	63
L3	7	5	2	2	84

Thirteen male monkeys (*Macaca fascicularis*) 5–8 years old, weighing 3.5–4.8 kg were studied. Four animals were naive controls. Nine received daily (0.35 mg per kg) injections of MPTP i.v. for 4 days. Three animals (M1–3) who received no chloroquine were the untreated controls. Six animals were pretreated with chloroquine (4 mg per kg) intramuscularly; three (S1–3) were pretreated for 12 days, and three (L1–3) for 24 days. All six pretreated animals continued to receive chloroquine injections during MPTP administration and for 10 days following MPTP exposure. The neurological examination evaluated the spontaneous movement, tremor, tone and deep tendon reflexes. A 0 score in each category reflects a total loss of spontaneous movement, maximum tremor, maximum increase in tone, or maximum hyperreflexia. Deep tendon reflexes examined were brachial radials, knee jerk, and ankle jerk. Muscle tone tests evaluated protraction-retraction, abduction-adduction and flexion-extension in both upper and lower extremities. Tremor was rated on severity and the number of extremities involved. Spontaneous movement was evaluated in the morning over a 30 min period. The rating scale consists of arbitrary units which are weighted to reflect the disturbances most prominent in MPTP treated monkeys. Control values are the maximum (normal) score for each element of the exam and all control animals fell within 10% of the control values. The sum of scores for the four elements are multiplied by 5.26 to provide a total score with a value of 100 for control animals.

* Chloroquine level was 36% of that in other monkeys monitored.

Table 2 Biochemical effects of chloroquine on MPTP neurotoxicity in monkeys

	Chloroquine in plasma (ng ml ⁻¹)	DA	HVA	Putamen HVA/DA	TH	DA	HVA	Caudate HVA/DA	TH
		(nmol per g tissue)	(nmol per g tissue)	(nmol per g tissue)	(% control)	(nmol per g tissue)	(nmol per g tissue)	(nmol per g tissue)	(% control)
Control									
C1		47.0	43.5	0.9		62.4	43.8	0.7	
C2		42.1	12.1	0.3		39.1	17.4	0.4	
C3		73.4	34.6	0.5		79.1	17.3	0.2	
C4		64.1	36.2	0.6		53.5	29.9	0.6	
MPTP									
M1		1.2	5.6	4.7	11.8	0.6	4.1	7.5	6.9
M2		1.4	2.6	1.8	10.8	0.6	1.7	2.8	8.4
M3		1.8	5.8	3.2	21.0	0.9	8.7	9.4	17.9
Short-term chloroquine and MPTP									
S1	ND	5.2	10.1	1.9	20.0	1.1	8.0	7.5	14.0
S2	ND	5.0	8.6	1.7	23.3	1.3	10.5	8.1	18.7
S3	300	11.2	9.8	0.9	29.8	7.6	9.2	1.2	22.6
Long-term chloroquine and MPTP									
L1	120	1.1	5.2	4.7	7.2	0.5	5.7	11.1	8.1
L2	310	16.0	20.1	1.3	58.5	11.0	15.7	1.4	39.5
L3	370	29.6	18.3	0.6	101.0	17.3	5.9	0.3	77.1

To determine dopamine and HVA levels, brain samples (10–20 mg wet weight) were homogenized in 300 µl of 0.4 M perchloric acid and centrifuged at 4 °C for 10 min at 1,000g. Aliquots of the supernatants were analysed directly by reverse-phase high performance liquid chromatography (HPLC) on an ODS-3 column (Whatman Chemical Separation) with a Pellosil C₈ guard column (Alltech Associates). The mobile phase for this system was acetate-phosphate/methanol (95:5) which included EDTA and sodium heptanesulphonate as an ion-pairing agent (Bioanalytical Systems, Inc.). Detection was done electrochemically on a glassy carbon electrode (Bioanalytical Systems Inc.) at an applied voltage of 0.65 V. Plasma levels of chloroquine were determined by HPLC with UV detection. Samples were deproteinized with 0.2 vol. of 25% trichloroacetic acid followed by centrifugation at 4 °C for 30 min at 1,000g. Supernatants were removed, lyophilized overnight and reconstituted with 80 µl of 0.1 M perchloric acid. Samples were analysed directly using reverse-phase HPLC with a Whatman ODS-3 column (Whatman) and Pellosil C₈ guard column (Alltech). The mobile phase for this system was 40% acetonitrile: 0.1 M sodium phosphate (pH 3.0) with 75 mM perchloric acid. Absorption was detected at 343 nm (ref. 24). TH activity was assayed by the tritium release method of Nagatsu²⁵ and Levine²⁶ employing modified reaction conditions of Coyle²⁷. Supernatant fluid (50 µl) from brain homogenate (1 g tissue in 20 vol 50 mM Tris, pH 7.4) were added to 7-ml glass scintillation vials containing 10 µl 1 M K₂ HPO₄ pH 5.5, 10 µl dihydropteridine reductase, 1,000 units of catalase, 10 µl 0.5 mM L-tyrosine, 5 µl NADH (7 mg ml⁻¹), 5 µl 6 DL-6-methyl-5,6,7,8-tetrahydropterine (2.8 mg ml⁻¹), 5 µl FeSO₄ (2.78 mg ml⁻¹) and 1 µC of ring labelled [³H]tyrosine. Mixtures were incubated 30 min at 37 °C and the reaction terminated by adding 50 µl 3 M Na₂CO₃, pH 11.6. Toluene/isoamyl alcohol scintillant (5 ml) was then added directly to the vial and the contents mixed for 10 s. The aqueous and organic phases were allowed to separate and the ³H₂O extracted into the organic phase determined. ND, not determined.

effects of MPTP and, in two of three animals, long-term treatment provided more pronounced protection. We were puzzled why one of the monkeys receiving long-term chloroquine treatment (L-1) was not protected against the effects of MPTP. Chloroquine in plasma was assayed in four of the monkeys immediately before administration of MPTP. Monkeys S-3, L-2 and L-3, which were protected against the effects of MPTP, had 300, 310 and 370 ng ml⁻¹ chloroquine respectively ($\approx 1 \mu\text{M}$ of which half is bound to plasma protein). In contrast, monkey L-1, which developed a parkinsonian syndrome, had a plasma level of 120 ng ml⁻¹. Presumably, the failure of drug protection resulted from the diminished availability of chloroquine in this monkey.

We propose that chloroquine partially protects monkeys against MPTP neurotoxicity by interfering with the binding of MPP⁺ to neuromelanin in nigral neurons. Alternatively, chloroquine might compete with MPP⁺ for the dopamine uptake system. However, in our assays of [³H]MPP⁺ uptake into striatal synaptosomes, chloroquine only weakly inhibited MPP⁺ uptake, with 8% inhibition at $1 \mu\text{M}$ and 40% at $10 \mu\text{M}$ —concentrations not likely under our experimental conditions^{15,16}. In our preliminary experiments in monkeys, chloroquine did not affect the accumulation of MPP⁺ in the striatum. Four monkeys were injected i.v. with 200 μCi (0.035 mg per kg) [³H]MPTP and killed 2 days later. Two of these monkeys received chloroquine (4 mg per kg) for 20 days before MPTP treatment, continuing until they were killed. MPP⁺ accumulation in the striatum was the same in chloroquine-pretreated and control groups. In these experiments it was not possible to assay the substantia nigra where chloroquine might be expected to influence MPP⁺ accumulation.

Monoamine oxidase (MAO) inhibitors prevent the conversion of MPTP to the active metabolite MPP⁺ and thus protect against toxicity¹⁷⁻¹⁹. MAO activity was measured as the ability of mouse brain mitochondrial preparations to catalyse the conversion of [³H]MPTP to [³H]MPP⁺ as monitored by thin layer chromatography. Chloroquine at levels up to $100 \mu\text{M}$ failed to reduce MAO activity.

Because rodent mesencephalic dopamine neurons lack neuromelanin, chloroquine should not influence their response to MPTP. We treated C-57 mice (groups of 6) with MPTP (20 mg per kg for 4 days) and confirmed a 70% fall in [³H]mazindol binding and TH activity⁵. Pretreatment with chloroquine in the same regimen as in monkeys, failed to alter these effects of MPTP. Accordingly, chloroquine's actions on non-melanin related systems shared by monkey and mouse, such as lysosomal function²⁰, are probably not responsible for the protection in monkeys.

Because chloroquine concentrates efficiently in the retinal melanin of the eye²¹ it should block ocular accumulation of [³H]MPTP. Confirming results of Lyden *et al.*²², we observed 50% more radiolabel in eyes of C-57 black mice than albino CFW mice (groups of 6) treated with [³H]MPTP (1 μCi per 0.1 mg per kg). Chloroquine pretreatment (4 mg per kg for 10 days) lowered [³H]MPTP-associated radioactivity in the eyes of pigmented mice to the level of albino mice, but failed to alter radioactivity levels in albino mice.

The partial protection of monkeys from MPTP neurotoxicity elicited by chloroquine, together with the high-affinity interactions of MPP⁺ with neuromelanin^{11,12}, indicates that destruction of dopamine neurons in the substantia nigra by exposure to low doses of MPTP is dependent upon interactions of MPP⁺ with neuromelanin. By inhibiting the binding of MPP⁺ to neuromelanin, chloroquine may reduce intraneuronal sequestration of MPP⁺, resulting in reduced toxicity to organelles such as mitochondria²³.

Some cases of idiopathic Parkinson's disease might result from an environmental toxin which, like MPTP, is selectively concentrated in dopamine neurons and associated with neuromelanin. Conceivably, chloroquine or a similar drug could

be used to retard the progression of Parkinson's disease in human patients.

We thank Dawn C. Dodson and Rita Hollingsworth for secretarial assistance, and Dawn Ledbetter, Amy Compton, Tammy Dellovade, David Benham, and Laine Brainard for technical assistance. This work was supported by USPHS grants; by a USPHS training grant to R.J.D. and a Research Scientist award to S.H.S.; and by the Bernard Weinberg Fund, the Robert L. and Clara G. Patterson Trust, the Claster family and The Gustavus and Louise Pfeiffer Foundation.

Received 24 October 1986; accepted 6 April 1987.

1. Davis, G. C. *et al. Psychiat. Res.* **1**, 249-254 (1979).
2. Langston, J. W., Ballard, P., Tetrud, J. W. & Irwin, I. *Science* **219**, 979-980 (1983).
3. Burns, R. S. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 4546-4550 (1983).
4. Langston, J. W., Forno, L. S., Rebert, C. S. & Irwin, I. *Brain Res.* **292**, 390-394 (1983).
5. Heikkilä, R. E., Hess, A. & Duvoisin, R. C. *Science* **224**, 1451-1453 (1984).
6. Hallman, H., Lange, J., Olson, L., Strombert, I. & Jonsson, G. *J. Neurochem.* **44**, 117-127 (1985).
7. Kitt, C. A., Cork, L. C., Eidelberg, E., Joh, T. H. & Price, D. L. *Neuroscience* **17**, 1089-1103 (1986).
8. Chiba, K., Trevor, A. & Castagnoli, N. Jr *Biochem. biophys. Res. Commun.* **120**, 574-578 (1984).
9. Javitch, J. A., D'Amato, R. J., Strittmatter, S. M. & Snyder, S. H. *Proc. natn. Acad. Sci. U.S.A.* **82**, 2173-2177 (1985).
10. Lyden, A., Bondesson, V., Larsson, B. S. & Lindquist, N. G. *Acta Pharmac. Tox.* **53**, 429-432 (1983).
11. D'Amato, R. J., Lipman, Z. P. & Snyder, S. H. *Science* **231**, 987-989 (1986).
12. D'Amato, R. J., Benham, D. & Snyder, S. H. *J. Neurochem.* **48**, 653-658 (1987).
13. Larson, B. & Tjälve, H. *Biochem. Pharmac.* **28**, 1181 (1979).
14. Schwartzman, R. & Alexander, G. *Brain Res.* **358**, 137-143 (1985).
15. Grundman, M., Mikulikova, I. & Urubolusky, P. *Archs int. Pharmacodyn.* **197**, 45-52 (1972).
16. McChesney, E. W., Shekosky, J. M. & Hernandez, P. H. *Biochem. Pharmac.* **16**, 2444-2447 (1967).
17. Heikkilä, R. E., Manzino, L., Cabbat, F. S. & Duvoisin, R. C. *Nature* **311**, 467-469 (1984).
18. Fuller, R. W. & Hemrock-Leucke, S. K. *Life Sci.* **37**, 1089-1096 (1985).
19. Langston, J. W., Irwin, I., Langston, E. B. & Forno, L. S. *Science* **225**, 1480-1482 (1984).
20. DeDuve, C. *et al. Biochem. Pharmac.* **23**, 2495-2530 (1974).
21. Lawwill, T., Appleton, B. & Altstatt, L. *Am. J. Ophthalm.* **65**, 530-532 (1968).
22. Lyden, A., Bondesson, V., Larsson, B. S., Lindquist, N. G. & Olsson, L. I. *Acta Pharmac. Tox.* **57**, 130-135 (1985).
23. Nicklas, W. J., Vyas, I. & Heikkilä, R. E. *Life Sci.* **36**, 2503-2508 (1985).
24. Bergqvist, Y. & Frisk-Holmberg, M. *J. Chromat.* **221**, 119-127 (1980).
25. Nagatsu, T., Levitt, M. & Udenfriend, S. *Analyt. Biochem.* **9**, 122-126 (1964).
26. Levine, R., Pollard, H. & Kuhn, D. *Analyt. Biochem.* **143**, 205-208 (1984).
27. Coyle, J. *Biochem. Pharmac.* **21**, 1935-1944 (1972).

No evidence for preservation of somatostatin-containing neurons after intrastriatal injections of quinolinic acid

Stephen W. Davies* & Peter J. Roberts†

Department of Physiology and Pharmacology, University of Southampton, Bassett Crescent East, Southampton SO9 3TU, UK

Intrastriatal injections of excitotoxic amino acids and their analogues (for example kainate and ibotenate) elicit a pattern of neuronal degeneration that is similar in many respects to that observed in Huntington's disease¹⁻⁷. In this disease there is a progressive degeneration of most types of intrinsic neuron but somatostatin and neuropeptide Y levels are increased 3-5-fold⁸⁻¹¹. This may be attributed to the selective preservation of a sub-class of striatal aspiny neurons, in which these two peptides are co-localized together with the enzyme NADPH-diaphorase (refs 11-15). Beal *et al.* reported recently that following intrastriatal injections of quinolinic acid in rats, medium-sized aspiny neurons were selectively preserved and they suggested that quinolinic acid which is found in human brain might cause the neuronal degeneration seen in Huntington's disease⁴. We have used immunocytochemical and enzyme histochemical techniques to examine this selective toxicity but find no evidence to support this finding. We conclude

* Present address: Department of Anatomy and Embryology, University College London, Gower Street, London WC1E 6BT, UK.

† To whom correspondence should be addressed.

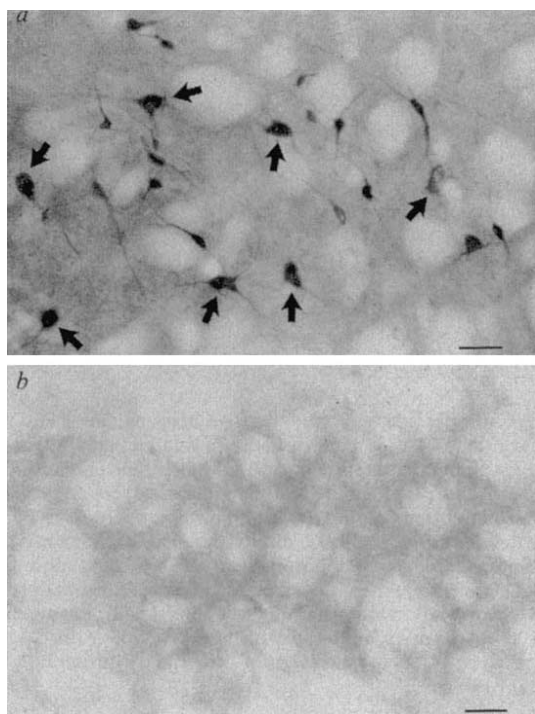


Fig. 1 Histochemical localization of AChE and diaphorase in *a*, uninjected striatum and *b*, striatum injected with quinolinic acid (60 nmol). Acetylcholinesterase-positive neurons are arrowed (calibration bar, 50 μ m). For histochemical detection of AChE, animals were pre-treated with di-isopropylfluorophosphonate (DFP, Aldrich, 2 mg per kg i.m.) in arachis oil, together with atropine (10 mg per kg i.p.). After 270 min, the animals were re-anaesthetized and perfused with 4% paraformaldehyde as previously described. Acetylcholinesterase activity was detected by the method of Karnovsky and Roots¹⁹, following detection of diaphorase activity according to the method of Vincent *et al.*²⁰.

that there are substantial differences between the immunocytochemical changes detected in postmortem Huntington's disease brain and those following quinolinic-acid-induced degeneration.

Twelve male Sprague-Dawley (175 g) and sixteen Wistar rats (175 g or 300 g) were anaesthetized with pentobarbitone (Sagatal; 60 mg per kg intraperitoneally (i.p.)) and positioned in a David Kopf small animal stereotaxic frame. A 30 gauge blunt-tipped Hamilton syringe needle was inserted into the striatum and quinolinic acid (2,3-pyridinedicarboxylic acid, Sigma) dissolved in 0.5 μ l phosphate-buffered saline (PBS), final pH 7.0, infused over a period of 5 min. The needle was left in place for a further 5 min and then withdrawn slowly. Two weeks after the injection, animals were re-anaesthetized, tracheotomized, artificially ventilated, and perfused transcardially through the ascending aorta with 50 ml warm (22 $^{\circ}$ C) heparinized saline, followed by either 300 ml of 4% paraformaldehyde in 0.1 M Sorensen's phosphate buffer, or a lysine-periodate/paraformaldehyde mixture, containing 2% paraformaldehyde¹⁶. The fixed brains were frozen with powdered dry ice, and cut on a sliding microtome in coronal section at 40 μ m. Every section throughout the striatum between the rostral pole of the nucleus accumbens, and the medial globus pallidus was collected. Sections were then processed either for the histochemical localization of acetylcholinesterase (AChE) and NADPH-diaphorase, or the immunocytochemical localization of somatostatin and neuropeptide Y. Alternate sections were stained with thionin.

Examination of thionin-stained sections two weeks after injection of 120 nmol quinolinic acid revealed a virtually complete loss of neuronal cell bodies in the lesioned area of the striatum,

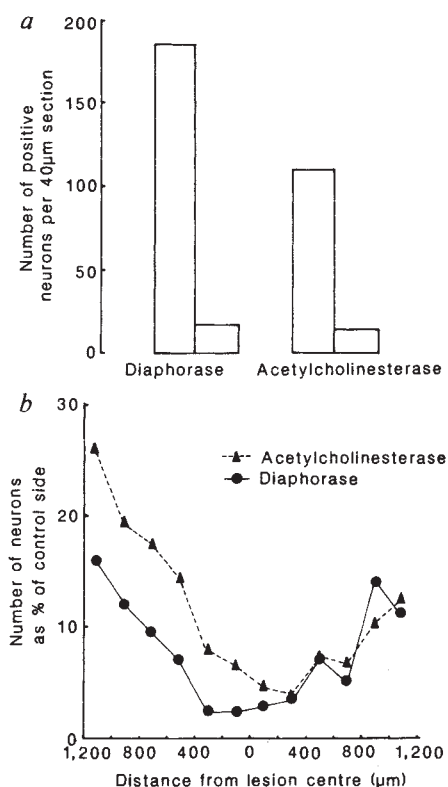


Fig. 2 *a*, Number of acetylcholinesterase- and diaphorase-positive neurons in control (left) and quinolinic acid (120 nmol) lesioned (right) striatum. Results are from four animals, where a 1:5 series of sections throughout the entire striatum was collected. Each column represents the mean number of positive cells in each striatum from 50 sections. *b*, Distribution of cell loss throughout the striatum, anterior and posterior to the injection site, for neurons positive for acetylcholinesterase (▲) and NADPH-diaphorase (●). Each point is the mean from four animals. There was no significant difference between the cell losses for both cell types at any position within the striatum.

with extensive gliosis. Peripheral to this area of damage, the tissue showed fairly normal morphology, and the head and tail of the striatum and the globus pallidus appeared morphologically normal. The contralateral side of the brain remained intact and showed normal morphology. Moreover, there was no sign of neuronal degeneration in the hippocampus or piriform cortex, areas that are particularly sensitive to the distant neuronal degeneration frequently found following intrastriatal kainic acid. The histochemical detection of AChE and NADPH-diaphorase revealed discrete neuronal perikarya (Fig. 1) which were differentiated easily, both on the basis of somal size (acetylcholinesterase-positive neurons: long axis, 31.54 ± 4.80 μ m; short axis, 17.28 ± 3.10 μ m; diaphorase-positive neurons: long axis, 18.37 ± 2.60 μ m; short axis, 8.84 ± 1.90 μ m) and colour (AChE-containing neurons are stained Hattchett's brown, whereas those containing NADPH-diaphorase are dark blue). The intrastriatal injection of 120 nmol of quinolinic acid completely eliminated both AChE- and diaphorase-containing perikarya throughout the entire lesion area as seen with nissl staining (Fig. 1). Quantification of this cell loss was achieved in two ways: Fig. 2*a* shows the mean number of surviving AChE and diaphorase neurons in alternate sections taken through the entire striatum. Cell counts (NADPH diaphorase and AChE) were also made at progressively larger distances both anterior and posterior to the centre of the injection site (Fig. 2*b*). In no instance was any selective preservation of diaphorase-positive neurons observed. These findings were consistent for differing age and strain of rat. Intrastriatal injection of 120 nmol quin-

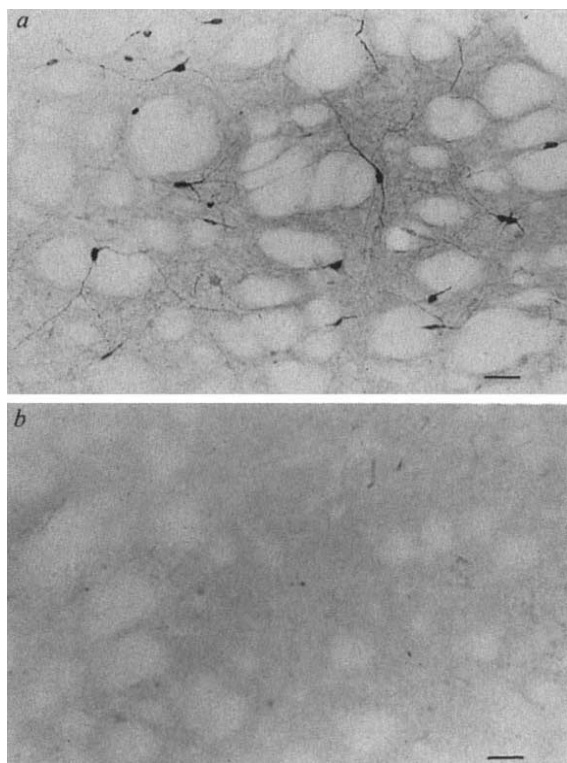


Fig. 3 Somatostatin-immunoreactive neurons, fibres and pre-terminal processes within *a*, control striatum and *b*, striatum lesioned by quinolinic acid (60 nmol). Photomicrographs are from identical parts of either striatum, on the same section. (Calibration bar, 50 μ m.)

Methods. Sections were incubated free-floating in rabbit anti-somatostatin (diluted 1:4,000 in PBS, containing 0.2% Triton X-100 and 1% normal goat serum) for no more than 4 days. The incubation was stopped by thorough washing of the sections in PBS and they were subsequently processed according to the avidin-biotin complex method of Hsu *et al.*²¹, using a commercially available ABC-kit (Vector Laboratories) as previously described²². The somatostatin antibody was raised in rabbits against synthetic somatostatin 14, linked to thyroglobulin with carbodiimide. All positive immunostaining was blocked by incubation of the antisera with 1 nmol ml⁻¹ somatostatin 14, or 10 nmol ml⁻¹ somatostatin 28, or following omission of the primary antisera.

olinic acid resulted in a comparable lesion in 175-g Wistar (46 days old), 175-g Sprague Dawley (48 days old) and 300-g Wistar (77 days old) rats. Injection of smaller amounts (60 or 30 nmol) of quinolinic acid resulted in a progressively smaller lesion core, seen in both nissl and AChE/diaphorase stained sections. Again, there was no indication of selective preservation of medium-sized diaphorase-positive neurons in comparison with the control tissue. Preparations stained immunocytochemically for either somatostatin (Fig. 3) or neuropeptide Y (Fig. 4) exhibited discrete perikarya with a dense staining of fibres and pre-terminal processes in the striatum. Injection of 60 nmol quinolinic acid resulted in an almost complete loss of neuronal cell bodies containing these two peptides, together with their associated processes. Neurons immunoreactive for these two peptides in adjacent structures (cortex, septum and hippocampus) appeared normal. The area in the striatum lacking somatostatin or neuropeptide-Y-containing neurons, following 60 nmol quinolinic acid, matched very closely the area from which AChE and diaphorase-positive neurons were absent (Figs 1, 3 and 4).

The close agreement between the results obtained with histological (thionin), histochemical (AChE and diaphorase) and immunocytochemical (somatostatin and neuropeptide Y) methods suggests that, contrary to the findings of Beal *et al.*⁴,

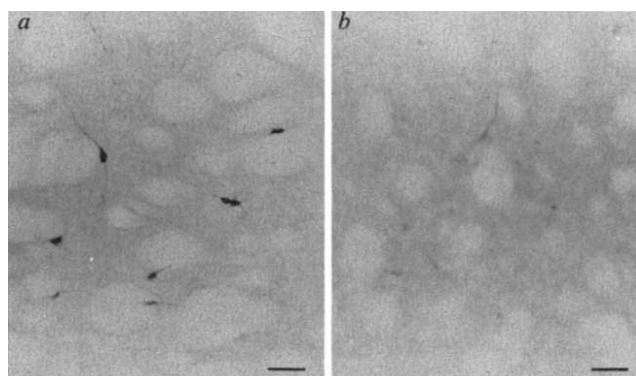


Fig. 4 Neuropeptide Y immunoreactive neurons, fibres and pre-terminal processes in *a*, control and *b*, quinolinic acid (60 nmol) lesioned striatum (calibration bar, 50 μ m). Neurons in *a* and *b* are from identical areas of either striatum in the same section. Sections were incubated free floating in rabbit anti-neuropeptide Y (Cambridge Research Biochemicals) diluted 1:4,000 in PBS containing 0.2% Triton X-100 and 1% normal goat serum for no more than 4 days. Sections were then processed as described for somatostatin (Fig. 3 legend). All immunostaining was abolished following incubation of the primary antibody with neuropeptide Y for 24 h at 4 °C, or following omission of the primary antisera.

somatostatin-containing neurons are not spared following intrastriatal injections of quinolinic acid. We are confident that this cell loss is specific and receptor mediated, because we were able to antagonize completely the lesion produced by 60 nmol quinolinic acid with an equimolar amount of (-)-2-amino-5-phosphono-valeric acid, a selective pharmacological antagonist of all the actions of quinolinic acid. The apparently normal levels of these peptides in the striatum when determined by radioimmunoassay, after intrastriatal injections of quinolinic acid, may be due to their presence in afferent fibres innervating the striatum. It is known that striatal levels of dopamine, noradrenaline and their metabolites increase substantially following quinolinic acid-induced lesions¹⁷. Previous studies have indicated that 25–40% of striatal somatostatin may be contained in striatal afferents¹⁸.

In conclusion, our study fails to provide any evidence to support the notion that the neurotoxic actions of quinolinic acid in rat striatum are akin to the neuropathological changes occurring in the Huntington's disease brain. In the quest for endogenous excitotoxic mechanisms underlying neurodegenerative disorders, it is perhaps unreasonable to equate acute neurotoxic changes induced experimentally over a period of days with those occurring in disease states over decades. Quinolinic acid may or may not be a prime candidate for a function in the aetiology of Huntington's disease. However, this will be determined only by the direct study of the neurochemistry of this compound in normal and Huntington's disease brain tissue.

This work was supported by a grant from the Wellcome Trust. We thank Dr Peter Keen, University of Bristol, for somatostatin antisera.

Received 18 September 1986; accepted 7 April 1987.

- Coyle, J. T. & Schwarcz, R. *Nature* **263**, 244–246 (1976).
- McGeer, E. G. & McGeer, P. L. *Nature* **263**, 517–519 (1976).
- Schwarcz, R. *et al. Expl Brain Res.* **37**, 199–216 (1979).
- Beal, M. F. *et al. Nature* **321**, 168–171 (1986).
- Köhler, C. & Schwarcz, R. *Neuroscience* **8**, 819–835 (1983).
- McBean, G. J. & Roberts, P. J. *J. Neurochem.* **44**, 247–254 (1985).
- Schwarcz, R., Whetsell, W. O. & Mangano, R. M. *Science* **219**, 316–319 (1983).
- Spokes, E. G. S. *Trends Neurosci.* **4**, 115–118 (1981).
- Aronin, N. *et al. Ann. Neurol.* **13**, 519–526 (1983).
- Beal, M. F., Bird, E. D., Langlais, P. J. & Martin, J. B. *Neurology* **34**, 663–666 (1984).
- Dawbarn, D., De Quidt, M. E. & Emson, P. C. *Brain Res.* **340**, 251–260 (1985).
- Ferrante, R. J., Kowall, N. W., Martin, J. B. & Richardson, E. P. Jr *J. Neuropath. exp. Neurol.* **4**, 325 (1985).
- Chesselet, M. F. & Graybiel, A. M. *Neuroscience* **17**, 547–573 (1986).
- Sandell, J. H., Graybiel, A. M. & Chesselet, M.-F. *J. comp. Neurol.* **243**, 326–334 (1986).

15. Vincent, S. R. *et al.* *J. comp. Neurol.* **217**, 252-263 (1983).
16. McLean, I. W. & Nakane, P. K. *J. Histochem. Cytochem.* **22**, 1077-1083 (1974).
17. Mazzari, S., Aldinio, C., Beccaro, M., Toffano, G. & Schwarcz, R. *Brain Res.* **380**, 309-317 (1986).
18. Beal, M. F. & Martin, J. B. *Brain Res.* **266**, 67-73 (1983).
19. Karnovsky, M. J. & Roots, L. *J. Histochem. Cytochem.* **12**, 219-221 (1964).
20. Vincent, S. R., Staines, W. A. & Fibiger, H. C. *Neurosci. Lett.* **35**, 111-114 (1983).
21. Hsu, S. M., Raine, L. & Fange, R. H. *J. Histochem. Cytochem.* **29**, 577-580 (1981).
22. Davies, S. W. & Köhler, C. *Anat. Embryol.* **173**, 45-52 (1985).

BEAL ET AL. REPLY—Davies and Roberts contend that quinolinic acid striatal lesions do not spare somatostatin (NADPH-diaphorase) neurons. We have attempted to resolve the discrepancy by carefully re-examining our own material. There is one major methodological difference between the two reports. Davies and Roberts examined regions containing the lesion cores in which there is almost total neuronal loss and marked gliosis as depicted in their Fig. 1 and the data given in Fig. 2a. They found no selective sparing of neurons within this region and our own findings are in agreement with that result. Since we believe that sparing of somatostatin neurons is relative and not absolute, we make our counts in regions posterior to the lesion, where there is a depletion of total Nissl-stained neurons in the range of 50%. Within this region we find significant increases in NADPH-diaphorase neurons relative to acetylcholinesterase neurons. We have further examined this issue by counting numbers of NADPH-diaphorase neurons relative to the total number of Nissl-stained neurons. NADPH-diaphorase neurons are significantly increased as a percentage of the total neuronal population in this area. These results are

consistent with the findings of Choi and colleagues who showed that quinolinic acid results in relative, but not absolute, sparing of NADPH-diaphorase neurons in cortical cell cultures (*Science* **234**, 73-76; 1986).

Davies and Roberts do not take issue with our neurochemical data. We have replicated these results in numerous subsequent experiments and they have been replicated by others (Nemeroff, Kitt, Dissette and Schwarcz, personal communication). Davies and Roberts explain these results by asserting that somatostatin may be contained in striatal afferents. However, this cannot explain differential effects of kainic acid and quinolinic acid on somatostatin levels, since both produce axon-sparing lesions.

It is likely that quinolinic acid lesions produce sparing of both acetylcholinesterase and NADPH-diaphorase neurons relative to other neuronal populations. We have recently found that acetylcholinesterase neurons are also spared in Huntington's disease (Ferrante *et al.*, *Brain Res.* **411**, 162-166, 1987). These findings taken as a whole strengthen the validity of our original observations that quinolinic acid lesions are an animal model of Huntington's disease.

M. F. BEAL
N. W. KOWALL
J. B. MARTIN

Department of Neurology,
Massachusetts General Hospital
and Harvard Medical School,
Boston,
Massachusetts 02114, USA

The third component of complement (C3) is responsible for the intracellular survival of *Leishmania major*

David M. Mosser & Paul J. Edelson

Department of Pediatrics, Division of Infectious Diseases and Immunology, New York Hospital and Cornell University Medical College, New York, New York, USA

Leishmania are obligate intracellular parasites of mononuclear phagocytes. We^{1,2} and others³ have shown that the promastigote form of all species of *leishmania* activates complement from non-immune serum and that this activation can result in parasite lysis. This work, as well as earlier *in vivo* studies⁴, suggested that complement is an important component of host defence against leishmaniasis. We now present evidence that parasite complement fixation, in addition to increasing parasite phagocytosis^{5,6}, is required for the intracellular survival of *leishmania* in macrophages. We specifically show a strong correlation between parasite C3 fixation and intracellular survival. We attribute this survival, in part, to a decrease in the magnitude of the macrophage respiratory burst which is triggered by complement-coated, as opposed to uncoated, parasites.

Labelled *Leishmania major* promastigotes are added to monolayers and the number of intracellular organisms is determined at either 24 or 48 h and compared to the number of parasites which were phagocytized during the first hour. The number of organisms associated with the monolayer at one hour is determined with a parasite radiobinding assay, which measures the total number of macrophage-associated organisms¹. In separate assays, using two independent methods, we have determined that more than 85% of organisms in all groups studied are internalized during this initial incubation period.

Over a wide range of parasite inputs, in the absence of serum, resident mouse macrophages kill over 95% of the ingested promastigotes, but when fresh non-immune serum is included,

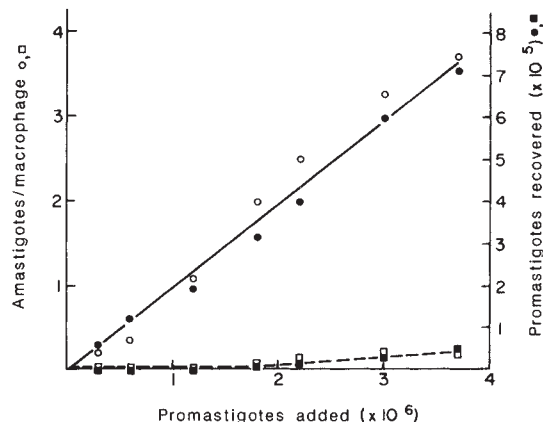


Fig. 1 The intracellular survival of *L. major* in resident murine macrophages. Increasing numbers of promastigotes were added in the presence (circles) or absence (squares) of 4% mouse C5D serum. Left axis, average number of morphologically intact amastigotes per macrophages counted on acridine orange-stained monolayers 48 h after infection (open symbols). The right axis indicates the number of viable organisms present per monolayer measured by the parasite reculture method which involves counting the number of promastigotes which transform from amastigotes as described⁷. Amastigotes are released from hypotonically lysed monolayers 24 h after infection (closed symbols). In separate experiments we have determined that the number of intracellular organisms remains constant from 24-48 h by this assay.

there is more than a 10-fold enhancement in survival (Fig. 1). Because of the increase in binding due to serum (opsonization) the average number of parasites bound per macrophage is higher, which might account for this increased survival. We therefore measured intracellular survival after the parasite input had been adjusted by two-thirds to three-quarters to achieve similar binding in the presence or absence of serum. Table 1 shows that ~95% of the 3.1 organisms bound per macrophage in the absence of serum are killed. In addition, fewer than 5%

of the macrophages contain organisms at 48 h (data not shown). In the presence of non-immune serum, macrophages which bind 4.1 parasites per cell support an average of 1.1 organisms/cell at 48 h. This enhancement does not depend on the continuous presence of fresh serum, as parasites which are briefly preopsonized with serum and then washed before their addition to monolayers survive as well as those ingested in the continuous presence of serum (Table 1 and Fig. 2). Heated serum, or serum treated with EDTA, does not support intracellular survival, suggesting a role for complement, but serum lacking C5 or C8 are fully competent, ruling out a requirement for the terminal complement components (Table 1). Normal human serum specifically depleted of C3 (C3D) was obtained by exposing serum to a monospecific antibody to C3 (Dako, California) for 1 h followed by antibody removal with Pansorbin (Calbiochem, California). C3-depleted serum cannot opsonize leishmania for enhanced parasite uptake and is unable to mediate intracellular survival. The addition of 250 U of purified C3 (Cordis, Miami, Florida) to this serum (C3D+C3) partially restores parasite intracellular survival (Table 1).

Initially, parasite opsonizations were performed for 7 min in serum which was deficient in one of the late complement components, to avoid parasite lysis. These opsonized parasites are fully viable and, when placed back into culture, show growth curves which are identical to untreated organisms. A slightly shorter exposure (5 min), also allows the use of intact human serum without parasite lysis. This brief preopsonization in intact serum results in the deposition of 4×10^4 molecules of C3 per parasite (determined using iodinated C3 as described previously⁷) and is sufficient to enhance parasite intracellular survival optimally. This suggests that complement fixation may be beneficial to the parasite *in vivo* even in normal hosts with an intact lytic complement pathway.

That this brief exposure to serum was sufficient to deposit C3b on the surface of the parasites (Table 1) was also shown by indirect immunofluorescence using a monoclonal antibody specific to C3b⁸ generously provided by P. J. Lachmann. All opsonizing conditions which allow enhanced parasite survival result in C3 deposition on the parasite surface (C8D, C5D, NHS, C3D+C3). Fluorescence is uniform throughout the parasite population with all parasites staining positive for bound C3. All conditions which do not result in C3 deposition fail to mediate intracellular survival (Table 1).

Because *Leishmania* are sensitive to superoxide and other toxic oxygen products^{9,10}, we compared the respiratory burst triggered by parasites in the presence or absence of serum. By measuring the respiratory burst during the phagocytosis of para-

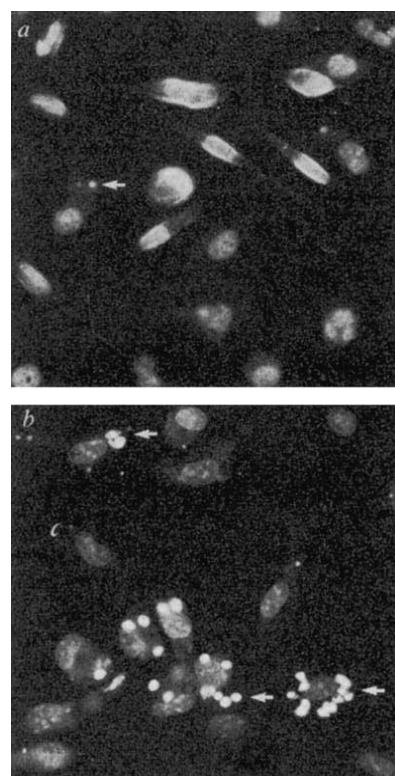


Fig. 2 Acridine orange (AO) staining of macrophage monolayers infected 48 h earlier with *Leishmania major* promastigotes which were preincubated with *a*, saline or *b*, C5-deficient (C5D) mouse serum. Parasites were resuspended to a concentration of $2 \times 10^8 \text{ ml}^{-1}$ in phagocytosis buffer² and exposed to either saline or 7% C5D for 7 min. Parasites were then washed twice and added to macrophages for 45 min at 36 °C. To achieve similar numbers of parasites ingested, 3×10^6 parasites were added in *a* and 1×10^6 organisms were added in *b*. Macrophages on coverslips were then thoroughly washed, placed in fresh wells, and washed again at 2 and 12 h. At 48 h cells were washed in Hanks balanced salt solution (HBSS), stained for 30 s in a 1% AO solution in water, diluted 1:50,000 in HBSS. Monolayers were then washed and fixed in 2% formalin and examined under a fluorescence photomicroscope. Arrows point to intact intracellular organisms. Minimal cytoplasmic staining causes some intracellular organisms to appear to be extracellular. Acridine orange staining is performed at 48 h because heavily infected macrophages sometimes require a full day to degrade killed parasites completely.

Table 1 The interaction of *Leishmania major* promastigotes with murine macrophages

	Phagocytosis*	n	Viability†	n	Morphology‡	n	C3 bound§
Serum independent	3.1 ± 0.29	8	0.19 ± 0.10	6	0.16 ± 0.05	6	—
4% Human C8D	4.1 ± 0.33	7	1.83 ± 0.34	3	1.10 ± 0.1	5	+
Preopsonized¶							
Mouse C5D	4.5 ± 0.71	5	1.55 ± 0.19	4	1.71 ± 0.27	3	+
Normal human	3.3 ± 0.51	5	ND		1.10 ± 0.31	2	+
HI—HS*	5.1 ± 1.9	4	0.15 ± 0.05	2	0.11 ± 0.10	3	—
HS—EDTA**	3.2 ± 0.35	3	0.20	1	0.09 ± 0.10	2	—
HS—C3D	2.9 ± 0.32	2	0.25 ± 0.07	2	0.20 ± 0.14	2	—
HS—C3D+C3††	2.6 ± 0.19	2	0.87 ± 0.31	2	0.63	1	+
Heat killed Ltm‡‡	3	2	0	1	0	2	

* Number of organisms phagocytosed at one hour, determined by the radiolabelled parasite binding assay, as described¹.

† Number of viable parasites present at 24 h, determined by the reculture method⁷.

‡ Average number of intact intracellular organisms per macrophage at 48 h, determined by the acridine orange technique¹².

§ The presence of bound C3 was determined by immunofluorescence microscopy as described⁵ using a monoclonal antibody to C3 (ref. 8).

|| Mean $\pm$ standard error.

¶ Parasites were preopsonized in 7% serum for 7 min with the exception of those exposed to normal human serum (NHS) which were preopsonized for 5 min.

* Human serum was heat-inactivated for 2 h at 56 °C.

** Parasites were exposed to human serum containing 9 mM EDTA.

†† C3-depleted serum to which 250 U of C3 (Cordis) was added.

‡‡ Promastigotes were treated at 60 °C for 30 min.

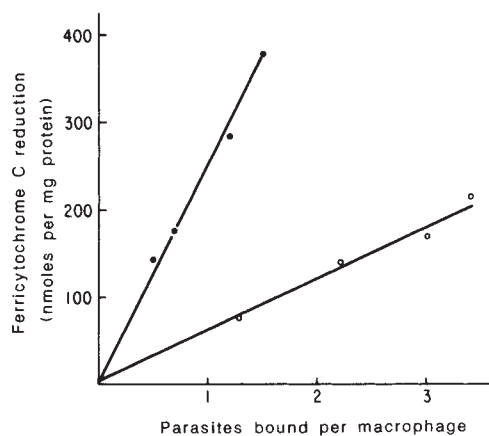


Fig. 3 The macrophage respiratory burst triggered during the phagocytosis of opsonized (○) or unopsonized (●) *Leishmania major* promastigotes, measured by the reduction of ferriochrome C, as described by Johnston¹³. Parasites were opsonized for 7 min in 7% human C8D. The ordinate is the number of nmoles of cytochrome C reduced h^{-1} per mg protein. The abscissa is the average number of radiolabelled organisms bound per macrophage, determined by the parasite radiobinding assay¹.

sites labelled with tritiated uracil, we could simultaneously measure the number of parasites bound and the respiratory burst they trigger on the same coverslip. Figure 3 shows that with either opsonized or unopsonized parasites, the magnitude of the respiratory burst increases in proportion to the number of parasites bound. C3-opsonized promastigotes, however, are poorer at stimulating a respiratory burst than are unopsonized organisms (Fig. 3). Equal numbers of parasites bound to the macrophage in the absence (2.2 ± 0.19 , $n=5$) or presence of serum (2.2 ± 0.29) trigger respiratory bursts of 178 ± 38 nmol O_2 per mg protein and 63 ± 12 respectively, a decrease of 65%. The respiratory burst triggered by opsonized zymosan, measured in parallel experiments, is 207 nmol O_2 per mg protein. The diminished production of superoxide measured in the presence of serum is not due to a non-specific protein quench as the burst is not diminished in the presence of 4% heat-inactivated serum.

We do not believe that serum immunoglobulins, which may be cross-reactive with the parasite² are responsible for enhanced intracellular survival. The brief time of preopsonization which we chose, 5–7 min, is not sufficient for the fixation of detectable amounts of immunoglobulin on the parasite's surface, by immunofluorescence microscopy (data not shown). Further, serum exposed to Protein A alone, performed as a control for the C3 depletion experiments, continues to mediate parasite survival. Finally, the addition of heat-inactivated antiserum to *L. major*, raised in rabbits, does not mediate parasite survival.

We therefore conclude that parasite-bound C3 plays a central role in the parasitism of macrophages by leishmania. Our previous work^{1,2} and that of others³ demonstrated that all species of leishmania activate complement. We have also shown that this activation results in an enhanced parasite uptake by macrophages via their complement receptors⁵. The present work demonstrates that even when equal numbers of *L. major* promastigotes are bound, the intracellular fate of C3-coated organisms is significantly different from that of non-opsonized organisms. Only those parasites with C3 on their surface establish a successful intracellular parasitism. We relate this successful parasitism, in part, to a decreased triggering of the respiratory burst by serum-opsonized organisms. Others have shown that macrophage complement receptors fail to trigger a respiratory burst¹¹.

This suggests that complement activation by leishmania and perhaps other organisms may be a common mechanism seized upon by intracellular parasites to ensure their successful entry into phagocytic cells which bear complement receptors.

This work was supported in part by the US National Institutes of Health.

Received 12 March; accepted 7 April 1987.

1. Mosser, D. M. & Edelson, P. J. *J. Immun.* **132**, 1501–1505 (1984).
2. Mosser, D. M., Burke, S. K., Coutavas, E. E., Wedgwood, J. F. & Edelson, P. J. *Exp. Parasitol.* **62**, 394–404 (1986).
3. Pearson, R. D. & Steigbigel, R. T. *J. Immun.* **125**, 2195–2201 (1980).
4. Poulter, L. W. *Clin. exp. Immun.* **40**, 25–35 (1980).
5. Mosser, D. M. & Edelson, P. J. *J. Immun.* **135**, 2785–2789 (1985).
6. Blackwell, J. M. *et al. J. exp. Med.* **162**, 324–331 (1985).
7. Mosser, D. M., Wedgwood, J. F. & Edelson, P. J. *J. Immun.* **134**, 4128–4131 (1985).
8. Lachmann, P. J., Pangburn, M. K. & Oldroyd, R. G. *J. exp. Med.* **156**, 205–216 (1982).
9. Murray, H. W. & Cartelli, D. M. *J. clin. Invest.* **72**, 32–44 (1983).
10. Pearson, R. D. & Steigbigel, R. T. *J. Immun.* **127**, 1438–1443 (1981).
11. Wright, S. D. & Silverstein, S. C. *J. exp. Med.* **158**, 2016–2023 (1983).
12. Pantasis, C. & Knicker, W. T. *J. Reticuloendothel. Soc.* **26**, 155–170 (1979).
13. Johnston, R. B., Godzik, C. A. & Cohn, Z. A. *J. exp. Med.* **148**, 115–127 (1978).

Genetic rescue of inviable hybrids between *Drosophila melanogaster* and its sibling species

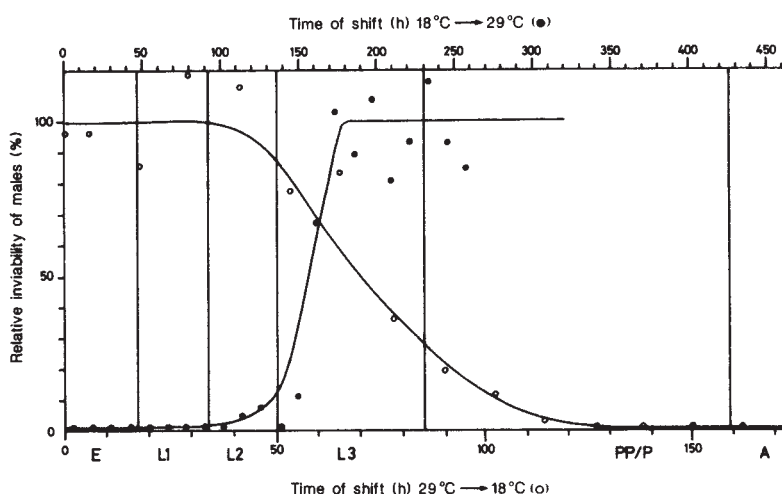
Pierre Hutter & Michael Ashburner

Department of Genetics, University of Cambridge, Cambridge CB2 3EH, UK

Post-mating mechanisms are central to the establishment of reproductive isolation between different, but closely related, species. Post-mating isolation mechanisms include hybrid breakdown, hybrid sterility and hybrid lethality^{1,2} and may, in some cases, be reinforced by pre-mating mechanisms such as ethological differentiation. In the *Drosophila melanogaster* species sub-group post-mating reproductive isolation is ensured by both the inviability and the sterility of hybrids³. For example when *D. melanogaster* females are crossed to *D. simulans* males the hybrid progeny are normally all female; the hybrid males die as third instar larvae^{4–6}. The viable hybrid females are totally sterile^{4,5}. Little is known of the genetic basis for either hybrid sterility or hybrid inviability, although Coyne and others^{7,8} have begun a genetic analysis of the sterility of hybrids within this species sub-group. We have discovered a single gene difference that rescues the otherwise inviable male hybrids from the cross between *D. melanogaster* females and males of its three closest relatives. The study of this locus may shed light on the genetic control of both speciation and development.

The *D. melanogaster* species sub-group consists of eight sibling species, six of which are endemic to the Afrotropical region or islands of the Indian Ocean. Four of these species are very closely related; they can only be reliably distinguished by their male genitalia and they will hybridize under laboratory conditions. These are the two cosmopolitan species *D. melanogaster* and *D. simulans*, *D. mauritiana*, endemic to the island of Mauritius, and *D. sechellia*, endemic to some islands of the Seychelles archipelago⁹. When *D. melanogaster* females are mated to males of the other species all the viable adult hybrids are sterile females. Males die as third instar larvae^{4,9,10}. Occasional surviving males are seen, these are X0 and result from the fertilization of nullo-X *D. melanogaster* eggs (the products of primary non-disjunction) by X-bearing sperm⁴. Watanabe¹¹ discovered an unusual strain (*K18*) of *D. simulans* from Japan that, when crossed to *D. melanogaster*, gave hybrid progeny of both sexes. The factor responsible for the rescue of the inviable hybrid class (*Lhr*) was mapped to chromosome arm 2R. Unfortunately, the paucity of genetic markers and aberrations in *D. simulans* has made it difficult to study this gene further. For this reason we screened 63 natural populations of *D. melanogaster* for an analogous mutation in this species. Females from these strains were crossed to males from a panel

Fig. 1 The relative viability of *Hmr/Y D. melanogaster/mauritiana* hybrid males in a temperature shift experiment. Eggs were collected for 12 h from *U79 D. melanogaster* ♀♀ × *D. mauritiana* ♂♂ and parallel cultures were maintained at either 18 °C (●) or 29 °C (○) for the times indicated and then shifted to the other temperature for the remainder of development. The relative viability of hybrid males is expressed as the (number of adult males)/(number of adult females) × 100. The approximate limits of the major developmental periods of the hybrid females are indicated: E, egg; L1-L3, larval instars; PP/P, prepupa/pupa; A, adult. In hybrid females pupariation occurred at ~228 h after the mid-point of the egg-laying period at 18 °C and at ~85 h at 29 °C.



of 26 *D. simulans* strains and the adult sex ratios of hybrid progeny, grown at 18 °C, were scored. It is known from previous work^{4,12,13} that the interspecific hybrids are generally more viable at a relatively low temperature. All but one of the crosses gave the expected result, namely sterile female hybrid progeny, with very few (and presumably non-disjunctive) males (Table 1). One strain of *D. melanogaster*, collected from Uman (Ukraine, USSR), consistently gave a relatively high proportion (about 5%) of hybrid males when crossed to different strains of *D. simulans*. This was not due to a high frequency of X-chromosome non-disjunction in the females of the Uman stock. By single-pair crosses within the Uman stock a sub-line (*U79*) was isolated which, when crossed to any strain of *D. simulans*, gave an almost 1:1 ratio of adult hybrid males to females (Table 1). We presume that the original population was polymorphic for a genetic factor responsible for rescuing the normally lethal hybrid males (which we call *Hmr*, for Hybrid male rescue). *D. melanogaster* females heterozygous for *U79* chromosomes and the dominantly marked balancer chromosomes *Basc*, *B*, *In(2L)Cy + In(2R)Cy*, *Cy* and *In(3LR)TM3, Ser* were crossed to *D. simulans* males. Surviving hybrid males were both *Cy* and *Cy*⁺ and both *Ser* and *Ser*⁺, however all were *B*⁺, showing that hybrid males survive only if they inherit the *U79* X chromosome (data not shown). A contribution of chromosome 4 to hybrid male survival was ruled out by a similar cross using *U79/ci*^D and *U79/ey*^D females.

Table 1 The outcome of crosses between females of the *U79* and other strains of *D. melanogaster* with males from 11 strains of *D. simulans*

	<i>D. melanogaster</i> ♀♀			
	<i>U79</i> ♂♂	<i>U79</i> ♀♀	Other strains ♂♂	Other strains ♀♀
<i>D. simulans</i> ♂♂				
Islamorada (Florida, USA)	136	143	0	2,480 (9)
Solway (Scotland)	208	217	1	843 (6)
Antibes (France)	212	219	0	3,609 (9)
Sevilla (Spain)	366	358	1	2,340 (9)
Dietikon (Switzerland)	262	270	2	4,513 (11)
Athens (Greece)	211	233	3	3,633 (18)
Nanyuki (Kenya)	302	313	2	3,094 (10)
St Louis (Missouri, USA)	313	308	0	773 (3)
Sao Paulo (Brazil)	117	118	0	1,362 (8)
Armidale (Australia)	204	201	0	1,039 (4)
Total	2,331	2,380	9	23,686
<i>K18</i> (Japan)	49	47	486	520 (8)

The numbers of *Hmr*⁺ strains of *D. melanogaster* crossed to each strain of *D. simulans* are indicated in brackets after the rightmost column. All cultures were grown at 18 °C. The *K18* strain of *D. simulans* carries the *Lhr* mutation (see text).

Table 2 Temperature-sensitive rescue of hybrid males by *Hmr* with *D. simulans*, *D. mauritiana* and *D. sechellia*

	Temperature (°C)	<i>D. melanogaster</i> ♀♀			
		<i>U79</i> ♂♂	<i>U79</i> ♀♀	Other strains ♂♂	Other strains ♀♀
<i>D. simulans</i> ♂♂	18	2,331	2,380	9	23,686
	25	14	305		
<i>D. mauritiana</i> ♂♂	18	854	823	0	2,586
	25	301	328		
	29	0	141		
<i>D. sechellia</i> ♂♂	18	274	692	0	167
	21	4	261		
	25	0	112		

Hmr is, therefore, an X-linked factor that acts zygotically to rescue the otherwise inviable *D. melanogaster/simulans* hybrid males. All these hybrids remain completely sterile. It has no obvious effect on the low viability of hybrid females from the reciprocal cross, that is of *D. simulans* females to *D. melanogaster* males (data not shown). Remarkably, *Hmr* also rescues the otherwise inviable hybrids between *D. melanogaster* females and either *D. mauritiana* or *D. sechellia* males (Table 2). These two species are very closely related to *D. simulans*: indeed hybrid females between either *D. mauritiana* or *D. sechellia* and *D. simulans* are fertile^{11,14,15}. The rescue of *D. melanogaster/simulans* and *D. melanogaster/mauritiana* hybrid males by *Hmr* is virtually complete at 18 °C. Rescue of *D. melanogaster/mauritiana* hybrids is also good at 25 °C but not at 29 °C. For the *D. melanogaster/simulans* hybrids rescue is even more temperature sensitive and is severely depressed at 25 °C. Hybrids between *D. melanogaster* and *D. sechellia* are most poorly rescued. Even at 18 °C only about one third of the expected males survive to adulthood (Table 2).

The temperature sensitivity of hybrid male rescue has been studied further by reciprocal temperature shifts between 18 °C and 29 °C of *D. melanogaster/mauritiana* hybrids (Fig. 1). Interpretation of these data is complicated by the fact that hybrid males are delayed in their development by 22–30 h at 18 °C and by 45–60 h at 29 °C. When these delays are taken into account, the data indicate a temperature-sensitive period for *Hmr* rescue in the second larval instar.

To locate *Hmr* more precisely, *Hmr/y v f D. melanogaster* females were crossed to *D. mauritiana* males. The data showed that hybrid males must carry the chromosome region from the *vermillion* region of the *Hmr* chromosome to survive. More accurate mapping was done with *Hmr/oc v* and *Hmr/lz v* females (Table 3). These data map *Hmr* to 1–31.84 (95% confidence limits 31.58–32.10), just distal to the *raspberry* gene. Cytogenetically this would indicate that *Hmr* is in polytene chromosome bands 9C1 to 9D3. The polytene chromosomes of

Table 3 Precise mapping of *Hmr* from crossing *Hmr/oc v* and *Hmr/lz v* *D. melanogaster* females to *D. mauritiana* males

<i>Hmr/oc v</i> ♀♀		<i>Hmr/lz v</i> ♀♀	
Phenotype	No.	Phenotype	No.
+	3,592	+	1,307
<i>oc</i>	330	<i>lz</i>	71
<i>v</i>	38	<i>v</i>	27
<i>oc v</i>	0	<i>lz v</i>	0

The numbers of adult males of each phenotype are shown. The standard map positions of the markers are *oc* 1-23.1, *lz* 1-27.7 and *v* 1-33.0.

U79 show no abnormality in this region (or elsewhere).

Morgan¹⁶ pointed out that *D. melanogaster/simulans* hybrids are inviable if they carry a Y chromosome from *D. simulans*. The Y chromosomes of *D. melanogaster* and *D. simulans* differ in two respects, the absence of a functional nucleolar organizer from the *D. simulans* Y and in their *Stellate* locus. The *Ste* gene¹⁷ affects the shape of the intra-nuclear crystals that are seen in primary spermatocytes of *D. melanogaster* X0 males. In this species this gene is repetitive on the X chromosome and similar sequences are also found on the Y chromosome. In *D. simulans* and *D. mauritiana* there are no X-linked copies of *Ste* and the number of *Ste*-related sequences on the Y chromosome is fewer than in *D. melanogaster*.

But the inviability of *D. melanogaster/simulans* hybrids cannot simply be due to the presence of a *D. simulans* Y chromosome, because genotypically identical female hybrids die if their mothers are *D. simulans* but live if they are *D. melanogaster*. This suggests that the mechanisms of hybrid death may differ between hybrid males and females and that there are interactions between the zygotic hybrid genome and maternal components⁴.

We have shown that a mutation of a single gene can rescue an otherwise inviable interspecific hybrid. Several questions remain to be answered. Among these are the nature of the *Hmr* mutation (for example, whether it is a novel gene function, or a gain or loss of function), the normal function of this gene in *D. melanogaster* (note that both *Hmr* and *Hmr*⁺ are fully viable) and the mechanism(s) of the hybrid rescue. The mechanism of rescue begs the question as to the genetic and developmental basis of hybrid lethality. It is unlikely that the *Hmr* gene has evolved simply as a genetic curiosity: it presumably has a function in the normal development of the species. It is also unlikely that *Hmr*⁺ has evolved to ensure the inviability of the species hybrids, because it is difficult to see how such an allele would become fixed within a species. We are now attempting to induce mutant alleles of *Hmr*, on both the *Hmr* and *Hmr*⁺ chromosomes, in order to study its function. Until such experiments are complete any suggestions as to its function could only be highly speculative.

P.H. was supported by a grant from the Fonds National Suisse de la Recherche Scientifique. We thank the following colleagues for sending strains: J. S. F. Barker, M. C. Carracedo, J. Coyne, D. L. Hartl, F. Lemeunier, P. A. Parsons, H. M. Robertson, R. Stocker, L. Tsacas, T. K. Watanabe, I. F. Zhimulev, B. W. van Ziyll and the Mid-American and Umeå Drosophila Stock Centres. J. Barrett gave statistical advice. We also thank Jerry Coyne, David Gubb and Trudy MacKay for discussions and comments on the manuscript of this paper.

Received 9 February; accepted 23 March 1987.

1. Dobzhansky, T. *Genetics of the Evolutionary Process* (Columbia University Press, 1970).
2. Mayr, E. *Animal Species and Evolution* (Belknap, Cambridge, Massachusetts, 1963).
3. Bock, I. *Evol. Biol.* **18**, 41-70 (1984).
4. Sturtevant, A. H. *Genetics* **5**, 488-500 (1920).
5. Sturtevant, A. H. *Carnegie Inst. Wash. Pubs.* **399**, 1-62 (1929).
6. Hadorn, E. *Rev. Suisse Zool.* **68**, 197-207 (1961).
7. Coyne, J. A. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4444-4447 (1984).
8. Coyne, J. A. & Charlesworth, B. *Heredity* **57**, 243-246 (1986).
9. Lemeunier, F., David, J. R., Tsacas, L. & Ashburner, M. in *The Genetics and Biology of Drosophila* Vol. 3e (eds Ashburner, M., Carson, H. L. & Thompson, J. N.) 147-256 (Academic, London, 1986).

10. Lachaise, D., David, J. R., Lemeunier, F., Tsacas, L. & Ashburner, M. *Evolution* **40**, 262-271 (1986).
11. Watanabe, T. K. *Jap. J. Genet.* **54**, 325-331 (1979).
12. Watanabe, T. K., Lee, W. H., Inoue, Y. & Kawanishi, M. *Jap. J. Genet.* **52**, 1-8 (1977).
13. Lee, W. H. *Jap. J. Genet.* **53**, 339-344 (1978).
14. David, J. R., Bocquet, C., Lemeunier, F. & Tsacas, L. *J. Genet.* **62**, 93-100 (1976).
15. David, J. R., Lemeunier, F., Bocquet, C. & Tsacas, L. *An. Génét.* **17**, 235-241 (1974).
16. Morgan, T. H. *Carnegie Inst. Wash. Pubs.* **399**, 201-222 (1929).
17. Livak, K. J. *Genetics* **107**, 611-634 (1984).

An amino-acid substitution involved in phenylketonuria is in linkage disequilibrium with DNA haplotype 2

Anthony G. DiLella, Joshua Marvit, Kelly Brayton* & Savio L. C. Woo

Howard Hughes Medical Institute, Department of Cell Biology and Institute of Molecular Genetics, Baylor College of Medicine, Houston, Texas 77030, USA

Phenylketonuria (PKU) is an autosomal recessive human genetic disorder caused by a deficiency of hepatic phenylalanine hydroxylase (PAH, phenylalanine 4-monooxygenase, EC 1.14.16.1). PKU is a common inborn error of amino-acid metabolism in caucasian populations and approximately 1 in 50 individuals are carriers of a PKU allele¹⁻². To define the molecular basis of PKU, we characterized twelve restriction fragment-length polymorphism (RFLP) haplotypes of the PAH locus in the northern European population and observed that 90% of the PKU alleles in this population are confined to four common RFLP haplotypes³⁻⁵. We have recently reported a splicing mutation in the PAH gene that is associated with RFLP haplotype 3 which is present at about 40% of mutant alleles⁶. We now report the molecular lesion associated with the RFLP haplotype 2 mutant allele. This defect is caused by a C-to-T transition in exon 12 resulting in an amino-acid substitution (Arg to Trp) at residue 408 of PAH. Direct hybridization analysis of the point mutation using a specific oligonucleotide probe demonstrated that this mutation is also in linkage disequilibrium with RFLP haplotype 2 alleles that make up about 20% of mutant PAH genes.

The human PAH gene contains 13 exons spanning 90 kilobases (kb) of DNA and codes for a protein of 451 amino acids⁷⁻⁸. Extensive genetic analysis of PKU families indicates that PAH deficiency in PKU patients is directly caused by mutations in the PAH gene⁹; RFLP haplotypes established at the PAH locus can be used to identify prevalent PKU alleles⁵. We recently reported a splicing mutation that is associated with mutant RFLP haplotype 3 which is the most frequent PKU allele among caucasians⁶. We wish to identify the remaining prevalent PKU mutations and determine whether they are also in linkage disequilibrium with their respective RFLP haplotypes. Twenty per cent of the PKU alleles analysed on 33 informative Danish PKU families are confined to haplotype 2 (Table 1), which constitutes only 5% of the normal PAH alleles in this population. A cosmid genomic library was constructed¹⁰ from leukocyte DNA isolated from a PKU individual homozygous for RFLP haplotype 2. The library was screened with the human PAH complementary DNA (hPAH247) probe and several corresponding genomic clones were isolated as previously described⁸. Sequence analysis of the exon containing regions of the mutant gene revealed a single mutation resulting from a C-to-T transition in the first base of codon 408, resulting in an amino-acid substitution (Arg → Trp) in exon 12 (Fig. 1a).

To ensure that the substitution at amino acid 408 was the direct cause of PKU, we performed site-directed mutagenesis using a specific oligonucleotide followed by expression analysis

* Present address: Purdue University, Department of Biochemistry, Life Science Research Building, West Lafayette, Indiana 47907, USA

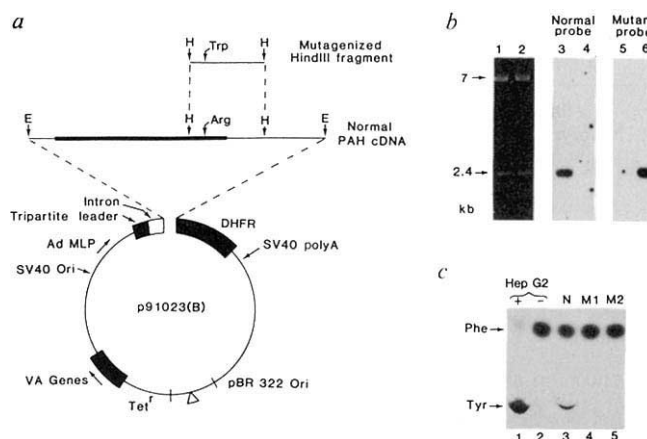


Fig. 2 Biochemical analysis of PAH. *a*, Cloning strategy. The internal *Hind*III fragment encoding the C-terminal 89 amino acids of PAH, the stop codon, and 316-bp of 3' noncoding sequence, was isolated from hPAH247 and inserted in the *Hind*III site of M13mp18 (ref. 11). A single-stranded recombinant M13 clone containing the insert in the sense-strand orientation was annealed to mutant-408 probe to generate the amino-acid 408 (Arg → Trp) substitution by site-directed mutagenesis, as described by Zoller and Smith¹². Point-mutant clones were identified by colony hybridization²³ using the mutagenic oligonucleotide as a probe. The mutagenized PAH *Hind*III fragment was then sequenced²² to ensure that only the oligonucleotide-directed mutation occurred. Replacement of the normal (with Arg 408) *Hind*III fragment of hPAH247 with the mutagenized (with Trp 408) *Hind*III fragment generated a full-length mutant PAH complementary DNA. Normal and mutant PAH cDNAs were then cloned into the *Eco*RI site of the eukaryotic expression vector p91023(B)¹³. *b*, Oligonucleotide specificity for the detection of the PKU mutation in cloned DNA by direct-gel hybridization. The *Eco*RI digests of the normal (lanes 1, 3 and 5) and mutant (lanes 2, 4 and 6) PAHcDNA/p91023B recombinant clones were electrophoresed on 1% agarose gels, stained with ethidium bromide (left panel) and processed for direct-gel hybridization as described previously^{6,14}. Dried gels were hybridized to oligonucleotide probes specific for normal (middle panel) and mutant (right panel) PAH cDNA sequences. *c*, PAH activity analysis. Normal and mutant PAHcDNA/p91023B recombinant clones were used to transfect monkey kidney (COS m6) cells by the suspension method of Chu and Sharp²⁴. PAH activity assays measuring the conversion of ¹⁴C-phenylalanine to ¹⁴C-tyrosine were as previously described¹⁵ on extracts prepared 36 h after transfection from ~10⁸ cells expressing normal (N) (lane 3) and mutant (M1 and M2) (lanes 4,5) PAH cDNA. PAH activity assays carried out on HepG2 extracts in the presence (+) (lane 1) and absence (–) (lane 2) of the synthetic cofactor 6-methyl-5,6,7,8-tetra-hydropterin served as the positive control.

The normal and mutant PAH cDNA/p91023(B) recombinant clones were used for transfer into monkey kidney COS cells for expression analysis. PAH enzyme assays were performed as previously described¹⁵ using protein extracts from cells transfected with the normal and mutant PAH cDNA. PAH activity was present in extracts from the human liver control (Fig. 2c, lane 1), and from cells transfected with the normal PAHcDNA/p91023B construct (Fig. 2c, lane 3). However, PAH activity was not detected in extracts from cells transfected with mutagenized PAHcDNA/p91023B constructs (Fig. 2c, lanes 4, 5), or in human liver extract in the absence of the synthetic cofactor 6-methyltetrahydropterin (Fig. 2c, lane 2). Whereas RNA analysis of transfected COS cells by Northern hybridization demonstrated that both the normal and mutant cDNA constructs expressed respective RNAs at similar levels, Western blot analysis of extracts from cells transfected with mutant PAH cDNA/p91023(B) failed to detect the presence of a mutated protein product (data not shown). The complete absence of PAH activity is indicative of the classical PKU phenotype¹ and can be attributed specifically to the single Arg-to-Trp substitution.

Fig. 1 Identification of a PKU amino-acid substitution. *a*, Sequence analysis of the region of exon 12 containing the codon for amino-acid 408. Cosmid library construction and gene characterization were carried out as described^{8,10}. The 5.4-kb genomic *Hind*III fragment (bottom) containing 151 base pairs (bp) of exon 11, the entire exon 12 sequence, and 383 bp of exon 13, was inserted into the *Hind*III site of M13mp18 (ref. 11). Dideoxynucleotide method sequence analysis²² of exon 12 was primed with a synthesized oligonucleotide (arrow) derived from the sequence data generated from the 3'-end of intron 11 of the normal PAH gene⁸. A comparison of the normal and mutant sequence illustrates the single base substitution. *b*, Oligonucleotide probes used to detect the substitution at amino-acid 408 in exon 12. *, The C-A mismatch between the normal probe and mutant gene sequence, and the mutant probe and normal gene sequence. Oligonucleotides were synthesized by an automated (Systec, Inc.) solid-phase phosphite triester method.

in cultured mammalian cells. Figure 1b shows the nucleotide sequence at the mutation site in the normal and mutant PAH genes. The mutant-408 probe is complementary to the sense strand of the normal exon sequence except for a C-A mismatch at the mutation site. The normal-408 probe is complementary to the antisense strand of the normal gene sequence and forms a C-A mismatch at the mutation site with the mutant gene. Because Arg 408 is within an internal 588-bp *Hind*III fragment of the normal human PAH cDNA clone pHPAH247, the *Hind*III fragment was subcloned into M13mp18 (ref. 11) and mutagenized¹² using the mutant-408 probe as the primer. The mutagenized *Hind*III fragment was used to replace the corresponding *Hind*III fragment in pHPAH247 to generate a full-length mutant PAH cDNA containing the Trp 408 mutation (Fig. 2a). The mutant and the normal PAH cDNAs were separately inserted into the eukaryotic high-expression vector p91023(B)¹³. The normal-408 and mutant-408 oligonucleotide probes were used for direct analysis of the mutation site in the normal and mutant PAHcDNA/p91023(B) recombinant clones as previously described¹⁴. Digestion of these clones with *Eco*RI generated a 2.4 kb fragment corresponding to normal (Fig. 2b, lane 1) and mutant (Fig. 2b, lane 2) PAH cDNA. Under stringent hybridization conditions, normal-408 probe hybridized only to the 2.4 kb *Eco*RI fragment of the normal cDNA (Fig. 2b, lane 3), and mutant-408 probe hybridized specifically to the mutant cDNA (Fig. 2b, lane 6). The data verified the presence of the transition sequence in the mutagenized PAH cDNA clone to be used for gene transfer experiments.

tion, because sequence analysis of the mutagenized *Hind*III fragment demonstrated that only the oligonucleotide-directed mutation had occurred (data not shown). Consequently, the C-to-T transition in exon 12 is indeed a PKU mutation.

The synthetic oligonucleotide probes were used under appropriate hybridization conditions to analyse the segregation of the normal and mutant PAH alleles in the family from which the mutant gene was isolated and characterized (Fig. 3). In this family, both parents carry a mutant haplotype 2 gene which hybridized to the mutant-408 probe (Fig. 3, lanes 1, 2). Because PKU is an autosomal recessive disorder, both parents are obligate carriers of the PKU trait and both must also contain a normal PAH allele. In this family, both normal alleles in the two parents are associated with haplotype 1, which hybridized to the normal-408 probe as expected (Fig. 3, lanes 5, 6). The proband in this family is homozygous for the mutant parental haplotype 2 alleles and DNA from the proband hybridized strongly to the mutant-408 probe (Fig. 3, lane 3) but not to normal-408 probe (Fig. 3, lane 7). The data confirmed that both mutant alleles in this family contain the same mutation. As expected, DNA from an unaffected sibling in this family who is heterozygous for mutant haplotype-2 and normal haplotype-1 alleles hybridized to both the mutant and normal-408 probes (Fig. 3, lanes 4 and 8).

We next determined the frequency of the mutation at amino-acid 408 in the Danish population and its degree of association with RFLP haplotype 2. RFLP haplotype analysis of 33 informative Danish PKU families provided a repertoire of mutant alleles for such an analysis. Genomic DNA samples isolated from PKU carriers and affected individuals of defined RFLP haplotypes were digested with *Pvu*II and analysed by hybridization using mutant-408 probe. For compound PKU heterozygotes bearing a mutant haplotype-2 allele, segregation analyses using parental DNA were performed. Table 1 summarizes the population genetic data. A total of 53% (35 out of 66) of the available PKU chromosomes and 27% (18 out of 66) of the available normal chromosomes were analysed. As shown in Table 1, all mutant haplotype-2 alleles hybridized with the mutant-408 probe whereas all the normal alleles did not, providing strong evidence that all mutant RFLP haplotype-2 alleles bear the same mutation. The lack of hybridization of the mutant-408 probe to three normal haplotype-2 alleles demonstrated that the mutation is not a constitutive part of the normal haplotype-2 genes. Also, the mutant-408 probe did not hybridize with any mutant alleles of other haplotypes (Table 1), suggesting that these alleles bear

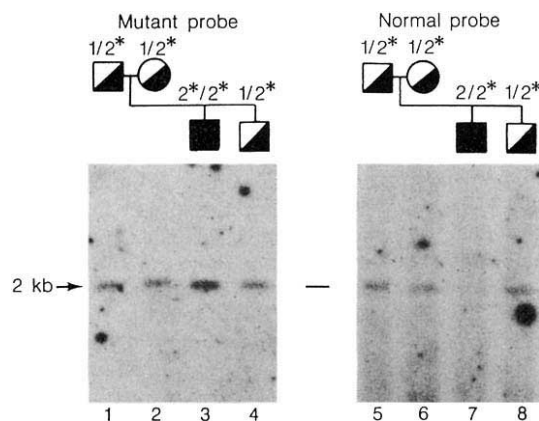


Fig. 3 Oligonucleotide hybridization analysis of a Danish PKU family. The probes used for lanes 1-4 and lanes 5-8 were the mutant and normal-408 oligonucleotides, respectively. Squares, male; circles, female; open symbols, normal haplotype; solid symbols, mutant haplotype. Lanes 1 and 5, father; lanes 2 and 6, mother; lanes 3 and 7, proband; lanes 4 and 8, unaffected child. DNA was isolated from leukocytes of family members, digested with *Pvu*II and analysed by dried-gel hybridization^{6,14}. Exon 12 is contained in a 2-kb *Pvu*II fragment⁶. The segregation of the PKU alleles (*) and the appropriate RFLP haplotypes of family members are shown at the top of each autoradiograph.

mutations different from the Arg-to-Trp substitution in exon 12. These results demonstrated a strong association between the mutation at amino-acid 408 and RFLP haplotype 2 in PKU.

Extensive RFLP haplotype analysis of the PAH locus in the Danish population demonstrated that 90% of the PKU genes are confined to haplotypes 1 to 4 (Table 1). We recently reported a splicing mutation in the PAH gene that is associated with RFLP haplotype 3, which constitutes about 40% of mutant alleles⁶. Our characterization of the haplotype-2 mutation corroborates our previous observations concerning strong association of PKU mutations with specific DNA haplotypes. To date, only sickle cell anaemia¹⁶ and β -thalassaemia mutations¹⁷ are known to be strongly associated with specific RFLP haplotypes. The linkage disequilibrium is presumably due to positive selection for the mutant β -globin alleles in malaria-infested region of the Mediterranean Basin. The high frequencies of the haplotype-2 and haplotype-3 PKU mutations observed here suggest that there may be a heterozygote advantage, as yet unidentified, for PKU. Alternatively, the splicing mutation may have occurred recently on a normal haplotype-3 chromosome and the codon-408 mutation may have occurred recently on a normal haplotype-2 chromosome, and both mutations could have spread in the population by founder effect. It has been proposed that the PKU allele is of Celtic origin and subsequently spread to various European populations¹⁸. Characterization of PKU mutations associated with the prevalent RFLP haplotypes throughout Europe will provide insight into the molecular evolution and origin of PKU mutations in man.

Although PKU is a heterogeneous metabolic disorder at the clinical level, it has been uncertain whether the observed variety of clinical phenotypes displayed by PKU patients resulted from multiple mutations in the PAH gene¹⁹. We recently reported that PKU patients who are either homozygotes or compound heterozygotes of the mutant haplotype-2 and haplotype-3 alleles follow a severe clinical course, whereas patients who carry mutant alleles of either haplotype 1 or haplotype 4 exhibit a milder clinical phenotype in general²⁰. Because the mutations in the PAH gene that are associated with haplotypes 2 and 3 result in negligible enzymatic activity, it is not surprising that patients who are either homozygous or heterozygous for these two mutant haplotypes follow a severe clinical course. Whether the clinical heterogeneity of PKU is due to combinations of different mutant PAH alleles will be resolved when the mutant

Table 1 Association of mutation of amino-acid 408 with mutant haplotype-2 alleles in Denmark

Haplotypes	Normal alleles		Mutant alleles		Mutant probe analysis†	
	Number*	%	Number*	%	Normal	PKU
1	23	34.8	12	18.2	0/11	0/9
2	3	4.6	13	19.7	0/3	9/9
3	2	3.0	25	37.9	0/2	0/8
4	21	31.8	9	13.6	0/1	0/4
5	7	10.6	0	0	0/1	0
6	0	0	2	3.0	0	0/1
7	7	10.6	1	1.5	ND	0/1
8	1	1.5	0	0	ND	0
9	0	0	1	1.5	0	ND
10	1	1.5	0	0	ND	0
11	1	1.5	1	1.5	ND	0/1
12	0	0	2	3.0	0	0/2

* Number of alleles observed in a total of 66 chromosomes analysed.

† Number of hybridizing alleles/number of genes analysed. All tested families were caucasian and had no history of consanguinity. Clinical criteria of the probands as affected PKU individuals and parents as heterozygotes in the families have been reported previously¹⁹.

ND, not determined owing to lack of DNA samples.

genes corresponding to haplotypes 1 and 4 are cloned and characterized²¹.

Collectively, haplotype-2 and haplotype-3 mutations constitute the majority of mutant alleles in the caucasian population of northern European ancestry (A.G.D., unpublished results). Characterization of PKU mutations in RFLP haplotypes 1 and 4 will clarify whether PKU is caused by a limited number of prevalent mutant alleles that recently spread throughout the caucasian race, or whether multiple PKU mutations arose independently in the population. If there are a limited number of prevalent PKU mutations, it will be possible to design a cassette of oligonucleotide probes to detect 90% of the PKU alleles in the caucasian population and thus detect carriers among random individuals without a previous family history of PKU.

We thank Dr Flemming Güttler of the Kennedy Institute in Denmark for the clinical samples used in this analysis, Mr David Nunn for technical assistance and Ms Kelly Porter for typing the manuscript. This work was supported by an NIH grant to S.L.C.W., who is also an Investigator of the Howard Hughes Medical Institute. A.G.D. is the recipient of a NIH postdoctoral fellowship.

Received 26 January; accepted 27 March 1987.

1. Sriver, C. R. & Clow, C. L. *New Engl. J. Med.* **303**, 1336-1343 (1980).
2. Sriver, C. R. & Clow, C. L. *A. Rev. Genet.* **14**, 179-202 (1980).
3. Woo, S. L. C., Lidsky, A. S., Güttler, F., Chandra, T. & Robson, K. J. H. *Nature* **306**, 151-155 (1983).
4. Lidsky, A. S. *et al. Am. J. hum. Genet.* **37**, 619-634 (1985).
5. Chakraborty, R. *et al. Hum. Genet.* (in the press).
6. DiLella, A. G., Marvit, J., Lidsky, A. S., Güttler, F. & Woo, S. L. C. *Nature* **322**, 799-803 (1986).
7. Kwok, S. C. M., Ledley, F. D., DiLella, A. G., Robson, K. J. H. & Woo, S. L. C. *Biochemistry* **24**, 556-561 (1985).
8. DiLella, A. G., Kwok, S. C. M., Ledley, F. D., Marvit, J. & Woo, S. L. C. *Biochemistry* **25**, 743-749 (1986).
9. Daiger, S. P. *et al. Lancet* **i**, 229-231 (1986).
10. DiLella, A. G. & Woo, S. L. C. *Meth. Enzym.* (in the press).
11. Messing, J. *Meth. Enzym.* **101**, 20-78 (1983).
12. Zoller, M. J. & Smith, M. *Meth. Enzym.* **68**, 389-395 (1979).
13. Wong, G. G. *et al. Science* **228**, 810-814 (1985).
14. DiLella, A. G. & Woo, S. L. C. *Meth. Enzym.* (in the press).
15. Ledley, F. D., Grenett, H. E., DiLella, A. G., Kwok, S. C. M. & Woo, S. L. C. *Science* **228**, 77-79 (1985).
16. Kan, Y. W. & Dozy, A. M. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5631-5635 (1978).
17. Orkin, S. H. & Kazanian, H. H. *A. Rev. Genet.* **18**, 131-172 (1984).
18. Woolf, L. I. *J. Ir. med. Ass.* **69**, 398-401 (1976).
19. Güttler, F. *Acta Paediatr. Scand. Stockholm Suppl.* **280**, 7-80 (1980).
20. Güttler, F. *et al. J. Pediatr.* **110**, 68-71 (1987).
21. DiLella, A. G. & Woo, S. L. C. *Molec. Biol. Med.* (in the press).
22. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
23. Benton, W. D. & Davis, R. W. *Science* **196**, 180-182 (1977).
24. Chu, G. & Sharp, P. A. *Gene* **13**, 197-202 (1981).

Use of restriction enzymes to detect potential gene sequences in mammalian DNA

Susan Lindsay & Adrian P. Bird

MRC Mammalian Genome Unit, King's Buildings,
West Mains Road, Edinburgh EH9 3JT, UK

Only a small proportion of the vertebrate genome codes for proteins. It would therefore be useful if genes, and in particular the sites at which transcription begins, could be identified in libraries of cloned DNA. Since many known vertebrate genes have distinctive sequences (HTF-islands) surrounding their transcription start sites¹, we wished to be able to select these sequences easily and to find out how diagnostic they are for genes. HTF-islands contain a high density of non-methylated CpG (ref. 2) and can be detected in chromosomal DNA as clustered sites for certain rare-cutting (C-G) restriction enzymes³. Identification of islands in chromosomal DNA is aided by methylation which blocks C-G enzyme sites in non-island DNA. This advantage is lost in cloned

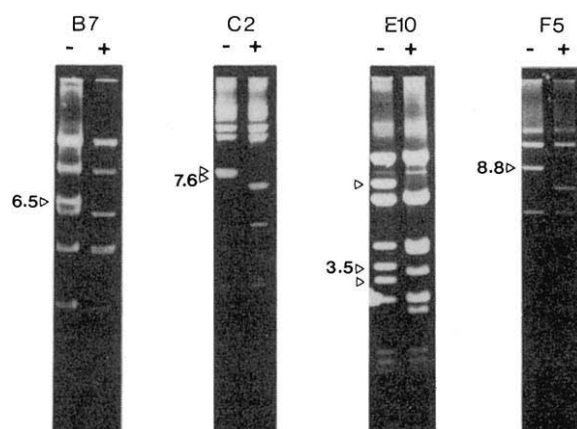


Fig. 1 *Hind*III (–) and *Hind*III/*Sac*II (+) digests of DNA from four cosmids with human DNA inserts derived from the X chromosome (B7, C2, E10 and F5). Arrows, *Hind*III fragments cut by *Sac*II. The sizes (in kb) are shown for *Hind*III fragments analysed subsequently (B7-6.5, C2-7.6, E10-3.5 and F5-8.8). The 8.8 kb fragment of F5 on later analysis proved to consist of the cosmid vector and 3.4 kb of insert DNA.

Methods. The library was constructed by the method of Ish-Horowitz and Burke¹². Right-hand and left-hand vector ends were prepared by appropriate digestion of the cloning vector pJB8. The Thy-B1-33-12 insert DNA was partially digested with *Mbo*I and size-selected (>23 kb and <50 kb) on a sucrose gradient. Thy-B1-33-12 is a mouse-human hybrid cell line with the X chromosome as the only detectable human chromosome⁴. After ligation of the vector ends to the insert DNA, the resulting cosmid clones were introduced into host bacteria Ed8767 and plated out onto nitrocellulose filters (Schleicher and Schuell) layered onto L-agar. Replica filters were taken and prepared for hybridization to nick-translated total human DNA¹³. Colonies on the master plate corresponding to those detected by the human DNA were picked, grown in L broth with ampicillin and cosmid DNA prepared¹². These DNAs were digested with *Hind*III, electrophoresed on a 0.8% agarose gel and bidirectionally blotted onto nitrocellulose membranes. One of the pair of filters was hybridized to nick-translated total human DNA and the other was hybridized to nick-translated total mouse DNA. Only those cosmids which hybridized to human DNA but not to mouse DNA were analysed further. DNA prepared from those cosmids was digested with *Hind*III or *Hind*III/*Sac*II according to the manufacturer's instructions (Anglian Biotechnology).

DNA, where CpG methylation is absent. We have calculated, however, that even in cloned DNA most sites for certain C-G enzymes should occur in HTF-islands. We tested this prediction using the enzyme *Sac*II and found that four out of four sites in separate cosmids from the human X chromosome were located in HTF-islands. Hybridization to Northern blots provided preliminary evidence that three of the islands were associated with genes.

Table 1 shows the calculated frequency of sites for various C-G enzymes in bulk DNA and in the HTF fraction (which is about 1% of genomic DNA). It is evident that sites for *Sac*II, *Bss*HII and *Eag*I, which have six-base-pair recognition sequences comprising only G and C and two CpGs per site, should be preferentially located in HTF-islands. Moreover, most islands are expected to contain sites for these enzymes (compare *Not*I sites which are found in only one-eighth of HTF-islands). As *Sac*II is a relatively cheap enzyme, we used it to screen for HTF-islands.

The cosmid library was constructed using DNA from the mouse-human hybrid cell line Thy-B1-33-12, which has the X chromosome as the only detectable human chromosome⁴. Thirty-two cosmids containing only human DNA were selected and *Hind*III and *Hind*III/*Sac*II digests performed. Four of the cosmids contained *Sac*II sites (B7, C2, E10 and F5; Fig. 1) and in two of these several *Hind*III fragments were cut (E10 and C2). The X-chromosome origin of the four cosmids was

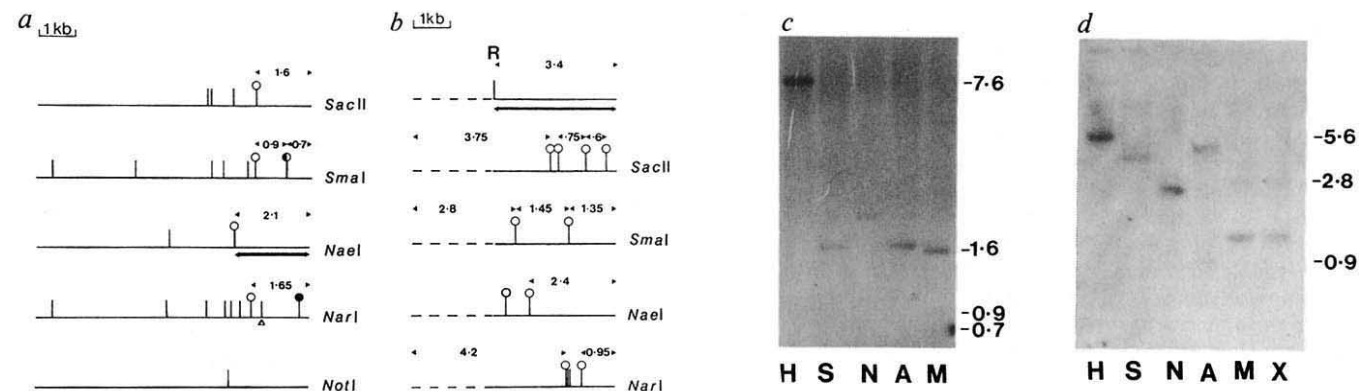


Fig. 2 C-G enzyme sites in cosmid subclones B7-6.5, C2-7.6, E10-3.5 and F5-8.8, and their methylation status in genomic DNA. *a, b*, Restriction maps of fragments C2-7.6 and F5-8.8 (respectively) for *SacII*, *SmaI*, *NaeI*, *NarI* and *NotI*. *R*, *EcoRI* site that marks the boundary between insert and cosmid vector in F5-8.8. The broken line to the left of *R* extends to a *HindIII* site that was identified in the genome—see *d*. Horizontal lines, *HindIII* fragments in the genome; vertical lines, restriction sites for the enzyme given at the end of each line. Fragments used as probes are illustrated by a thick black line under the relevant restriction map. Above each map are the sizes of fragments detected by the probes when hybridized to digests of sperm DNA. Open circle, unmethylated site; partially or fully shaded circle, partially or fully methylated site. *c, d*, Sample autoradiographs from hybridization of the 2.2-kb *HindIII*-*NaeI* fragment of C2-7.6 (*c*) and the 3.4-kb *EcoRI*-*HindIII* fragment of F5-8.8 (*d*) to sperm DNA digested with *HindIII* (H) or doubly digested with *HindIII* and *SacII* (S), *NaeI* (N), *NarI* (A), *SmaI* (M) or *XmaI* (X). *XmaI* is a methyl-insensitive isoschizomer of *SmaI*. *e*, Summary restriction and methylation maps. Horizontal lines, *HindIII* fragments as in *a* and *b*; vertical lines, C-G enzymes sites including *SacII* (s). The methylation status of the sites is indicated as above. The maps for B7-6.5, C2-7.6 and E10-3.5 were each constructed using several different probes (in *a* and *c*, the results using only one of the C2-7.6 probes are shown). We experienced difficulties in cutting some of the *NaeI* and *NarI* sites (open arrowheads) and those same sites also proved difficult to cut in the corresponding sub-clones. According to the manufacturer (New England Biolabs), *NaeI* and *NarI* have site-specific rates of cutting, that are independent of methylation, so partial cleavage does not necessarily indicate methylation. Only two sites showed reproducible differences in their methylation in sperm and leukocyte DNA. Site 1 (C2-7.6) is a

NaeI site which is fully methylated in leukocyte DNA; site 2 (C2-7.6) is a *SmaI* site which is unmethylated in leukocyte DNA. The map of clone C2-7.6 suggests that these differences occur at the edges of the HTF-island. Fragments used as probes to Northern blots are indicated by thick lines under the relevant clones. Where no probe is indicated the whole clone was used.

Methods. Human sperm DNA was prepared using standard procedures and digested with restriction enzymes according to manufacturers instructions. *NaeI*, *NarI* and *XmaI* were obtained from New England Biolabs and *SmaI* was obtained from Boehringer Mannheim. The digested DNAs were then electrophoresed on a 0.8 per cent agarose gel which was blotted¹⁴ to either nitrocellulose (Schleicher and Schuell) or Zeta probe¹⁵ (Biorad) filters. Probes were prepared by oligolabelling of denatured DNA fragments using random primers¹⁶. After the labelling step the probes were pre-competed in a volume of 100 μ l of 0.75 M NaCl, 0.075 M Na citrate with 1 mg of sonicated human placental DNA for 10 min at 68 $^{\circ}$ C to remove repeated sequences¹⁷. Hybridizations were carried out under standard conditions (to nitrocellulose filters) or in 225 mM NaCl, 22.5 mM Na citrate, 1% SDS, 0.5% dried milk powder and 0.5 mg ml⁻¹ denatured, sonicated salmon sperm DNA (to Zetaprobe filters)¹⁵.

Table 1 Calculated distribution of CpG endonuclease sites in mammalian DNA

Enzyme	Site	CpGs	Sensitivity to 5 mC	Sites per genome ($\times 10^{-4}$)		Percentage of sites in islands	Sites per island
				Bulk DNA	Islands		
<i>NotI</i>	GCGGCCGC	2	+	0.05	0.36	89	0.12
<i>EagI</i>	CGGCCG						
<i>SacII</i>	CCGCGG	2	+	1.2	3.5	74	1.2
<i>BssHII</i>	GCGCGC						
<i>SmaI</i>	CCCGGG						
<i>NaeI</i>	GCCGGC	1	+	4.8	3.5	42	1.2
<i>NarI</i>	GGCGCC						
(<i>NunI</i>)							
<i>MluI</i>	ACGCGT	2	+	2.7	1.0	27	0.34
<i>PvuI</i>	CGATCG						
<i>HpaII</i>	CCGG	1	+	120.0	33.0	21	11.1
<i>HhaI</i>	GCGC						

For the calculation of number of sites per genome, the base composition of bulk DNA was taken as 40% G+C, the CpG deficiency as 25% of the expected level and the genome size as 3×10^9 base pairs. The number of sites was calculated from the product of frequencies for each nucleotide in the fraction concerned, taking account of the CpG deficiency in the case of bulk DNA. HTF-islands were assumed to make up 1% of genomic DNA, to have a base composition of 65% G+C, and to show no deficiency of CpG. A range of sizes (usually 0.5 kb–3.0 kb) for HTF-islands has been found, so for the purpose of calculation an average length of 1,000 bp has been assumed. 5mC, 5-methylcytosine.

confirmed by hybridization to DNA from several mouse-human hybrid cell lines (data not shown).

If *Sac*II sites indicate the position of HTF-islands, then they should be within a cluster of other C-G restriction enzyme sites and the clustered sites should be non-methylated in DNA of germline and somatic tissues. To test this, restriction and methylation maps of four *Sac*II site-containing fragments (B7-6.5, C2-7.6, E10-3.5 and F5-8.8) were constructed using the methyl-sensitive C-G enzymes *Sac*II, *Nae*I, *Nar*I and *Sma*I (see Table 1). Only male genomic DNA was used to construct the methylation maps, as there is evidence that HTF-islands on the inactive X chromosome can be methylated⁵⁻⁷. Sample blots and detailed maps are shown for two of the fragments (C2-7.6 and F5-8.8) in Fig. 2, *a-d*. The methylation and restriction maps for all four fragments are summarized in Fig. 2*e*. It is evident that each fragment contains a cluster of sites for these enzymes which is non-methylated in DNA from sperm and leukocytes. The size of the clusters is typical for known HTF-islands (~1-2 kilobases (kb)).

If most HTF-islands are associated with genes¹, then a high proportion of our randomly-selected islands ought to detect transcripts when used as probes against Northern blots of poly(A)⁺ RNA. Hybridizations to HeLa-cell RNA were done under conditions where repeat sequences are competitively removed (see Fig. 2 legend). Using poly(A)⁻ RNA as a control is particularly important because the high G+C content of HTF-islands makes cross-hybridization to ribosomal RNA a frequent occurrence. Clear transcripts (bands present in the poly(A)⁺ but not the poly(A)⁻ tracks) were seen with three of the clones: B7-6.5 detected a single transcript of about 2 kb; E10-3.5 detected several transcripts of between 1.8 kb and more than 10 kb; and F5-8.8 detected a small transcript (0.2-0.5 kb) which was not accurately sized. C2-7.6 detected bands of similar size in poly(A)⁺ and poly(A)⁻ tracks indicating cross-hybridization to rRNAs. The signal was reproducibly more intense in the poly(A)⁺ track, however, so C2-7.6 may be detecting genuine transcripts of equivalent size to rRNAs.

Out of thirty-two starting cosmids, we detected four which had *Sac*II sites. We characterized one *Sac*II site-containing fragment from each cosmid and showed that all four fragments contained sequences characteristic of HTF-islands. This result is compatible with the theoretical expectation that three out of four *Sac*II sites would denote islands (see Table 1). The four islands were found in ~1,000 kb of X-chromosome DNA. Other uncharacterized *Sac*II sites in the cosmids (see cosmids E10 and C2, Fig. 1) may mark several more islands. In cosmid E10, for example, we have preliminary evidence that a second HTF-island occurs only 6 kb from the one in E10-3.5. Recently, several HTF-islands have been shown to occur in close proximity to each other at known gene loci both on the X chromosome⁸ and on an autosome⁹. It is possible that clustering of HTF-islands often occurs and may be an important feature of their organization. Three of the HTF-islands detected unequivocal transcripts in polyadenylated RNA from HeLa cells. We do not know the distribution of the transcripts in other human cell types nor how they map onto the island-containing fragments that detect them. Our results, however, support the prediction that most HTF-islands are associated with genes. Further evidence for this comes from the analysis of a randomly selected HTF-island from mouse DNA, which was found to be located at the 5' end of a pair of divergent messenger RNA-like transcripts (P. Lavia, D. Macleod and A. P. B., manuscript submitted).

In using this method to search for genes there are several points to consider. Firstly, variability of island sequences means that some islands lack *Sac*II sites and thus would be missed. The problem could be considerably reduced by simultaneously using a second C-G enzyme (such as *Bss*HII). Secondly, if an HTF-island does not detect transcripts this may either mean that it is not associated with a gene or that the gene is tissue-specific. Poly(A)⁺ RNA from a range of tissues may have to be

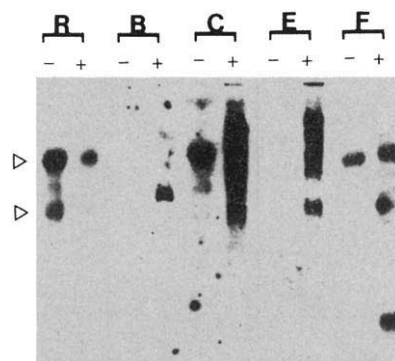


Fig. 3 Hybridization of HTF-island containing fragments to poly(A)⁺(+) and poly(A)⁻(-) RNA from HeLa cells. Probes were ribosomal RNA plasmid pXlr101 (R), B7-6.5 (B), C2-7.6 (C), E10-3.5 (E) and F5-8.8 (F).

Methods. RNA was extracted¹⁸ from HeLa-cell pellets by homogenizing once in 8 M guanidine HCl, 50 mM Tris and 10 mM EDTA (pH 7.5) and subsequently homogenizing twice in the above solution with 6 M guanidine-HCl instead of 8 M. The RNA was precipitated with a half volume of absolute ethanol after each homogenization and finally resuspended in a small volume of water to which ethanol was added to 70% for storage at -20 °C. Poly(A)⁺ RNA was prepared by passing total RNA down an oligo-dT cellulose column¹⁹. About 10 µg of poly(A)⁺ or poly(A)⁻ RNA were loaded per slot onto a 1.2% agarose gel in MOPS buffer (0.02 M MOPS, 0.005 M sodium acetate, pH 7.0, 0.001 M EDTA) and 6.8% formaldehyde. RNA was electrophoresed for 3-4 h at 100 V and the RNA then blotted to a Hybond-N (Amersham) filter. Probes were prepared either by oligolabelling¹⁶ or by nick translation. In all cases, the probes were pre-competed with sonicated human DNA (Fig. 2 legend). Hybridizations were carried out in 0.5 M sodium phosphate (pH 7.2), 7% SDS and 1 mM EDTA and the filters washed in 40 mM sodium phosphate (pH 7.2) and 1% SDS²⁰. Clone pXlr101 was a gift from R. Reeder (Hutchinson Cancer Centre).

probed to improve the chance of identifying a tissue-specific RNA. Thirdly, although it is likely that the majority of mammalian genes will be found to have HTF-islands¹ (they are associated with all polymerase II housekeeping genes characterized to date and several tissue-specific genes such as α -globin⁹ and Thy-1 (ref. 10)), many tissue-specific genes do not have HTF-islands (for example β -globin¹¹) and such genes will not be detected by the present method.

We thank Donald Macleod for assistance with RNA gels and blots, Patrizia Lavia for advice on RNA technology, Mary Taggart for help with clone F5-8.8, William Brown for constructing the cosmid library and Anne Deane for typing.

Received 17 February; accepted 9 April 1987.

1. Bird, A. P. *Nature* **321**, 209-213 (1986).
2. Bird, A. P., Taggart, M., Frommer, M., Miller, O. J. & Macleod, D. *Cell* **40**, 91-99 (1985).
3. Brown, W. R. A. & Bird, A. P. *Nature* **323**, 477-481 (1986).
4. Lund, E. *et al. Molec. cell Biol.* **3**, 2211-2220 (1983).
5. Yen, P. H., Patel, P., Chinault, A. C., Mohandas, T. & Shapiro, L. J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1759-1763 (1984).
6. Wolf, S. F., Jolly, D. J., Lunnen, K. D., Friedman, T. & Migeon, B. R. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2806-2810 (1984).
7. Wolf, S. F. *et al. Nucleic Acids Res.* **12**, 9333-9348 (1984).
8. Toniolo, D. *et al. EMBO J.* **3**, 1987-1995 (1984).
9. Bird, A. P., Taggart, M. H., Nicholls, R. D. & Riggs, D. R. *EMBO J.* **6**, 999-1004 (1987).
10. Kolsto, A.-B. *et al. Nucleic Acids Res.* **14**, 9667-9678 (1986).
11. Van der Ploeg, L. H. T. & Flavell, R. A. *Cell* **19**, 947-958 (1980).
12. Ish-Horowicz, D. & Burke, J. F. *Nucleic Acids Res.* **9**, 2989-2998 (1981).
13. Maniatis, T., Fritsch, E. F. & Sambrook, J. in *Molecular Cloning, a Laboratory Manual*, 304-305; 314 (Cold Spring Harbor Laboratory, New York, 1982).
14. Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
15. Reed, K. C. & Mann, D. A. *Nucleic Acids Res.* **13**, 7207-7221 (1985).
16. Feinberg, A. & Vogelstein, B. *Analyt. Biochem.* **137**, 266-267 (1984).
17. Sealey, P. G., Whittaker, P. A. & Southern, E. M. *Nucleic Acids Res.* **13**, 1905-1922 (1985).
18. Cox, R. A. *Meth. Enzym.* **12**, 120-129 (1968).
19. Maniatis, T., Fritsch, E. F. & Sambrook, J. in *Molecular Cloning, a Laboratory Manual* 197-198 (Cold Spring Harbor Laboratory, New York, 1982).
20. Church, G. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **18**, 1991-1995 (1984).

The solution structure of human epidermal growth factor

Robert M. Cooke, Anthony J. Wilkinson*,
Martin Baron, Annalisa Pastore, Michael J. Tappin,
Iain D. Campbell†, Harry Gregory* & Brian Sheard*

Department of Biochemistry, University of Oxford,
South Parks Road, Oxford, OX1 3QU, UK

* ICI Pharmaceuticals Division, Mereside, Alderley Park,
Macclesfield, SK10 4TG, UK

† To whom correspondence should be addressed

The epidermal growth factors (EGFs)¹⁻³ are powerful mitogens for a wide variety of cells in culture⁴; human EGF (hEGF), known as urogastrone, also inhibits gastric acid secretion *in vivo*¹. The transforming growth factors (TGF- α) are related to the EGF family both in sequence⁵ and activity^{6,7} and EGF-like sequences are often observed in a wide range of functionally unrelated proteins⁸. Attempts to examine the structure of EGF by diffraction methods have not yet succeeded because of difficulties with crystallization. We report here a three-dimensional structure of a biologically active derivative (residues 1-48) of the 53-residue human EGF. An analysis of high resolution ¹H nuclear magnetic resonance (NMR) spectra was used together with a combination of distance geometry, restrained energy minimization and restrained molecular dynamics methods. The three-dimensional structure provides a basis for understanding the properties of EGFs and for predicting the structures of homologous sequences in other proteins.

In a previous report on the secondary structure of hEGF⁹ we showed that residues 18-23 were paired with residues 34-28 in an antiparallel β -sheet. A type II β -turn was also identified between residues 34 and 37. We have since found a small antiparallel β -sheet-like structure linking residues 37-38 with 45-44. Weak nuclear Overhauser enhancements (NOEs) are observed between the α CH resonances of residues 3 and 23 and between those of residues 5 and 21, suggesting the N terminus may form the third strand of a triple-stranded β -sheet. A recent NMR study found similar secondary structural features in mouse EGF¹⁰, although the alignment of the N terminus with the other strands of the β -sheet is shifted by one residue.

A complete 3-D structure determination from NMR data requires the use of computer-based methods^{11,12}. The hEGF structures described here were initially produced using the distance geometry method¹³ with the program DISGEO. Refinement of these structures was accomplished by energy minimization and molecular dynamics¹⁴ using the GROMOS program. The major NMR-derived input for structure determinations arises from NOEs between resonances, which identify protons in close spatial proximity. Of the 186 NOE-derived distances used in these calculations, 46 were between resonances of the same

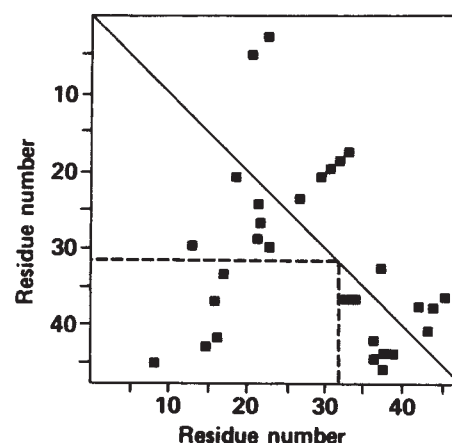


Fig. 1 NOEs observed between resonances from residues separated by more than one in the sequence. NOEs between backbone resonances are shown above the diagonal, whereas NOEs between side chain resonances and between backbone and side chain resonances are shown below the diagonal. The NOEs included in the dashed box are those which occur between the two domains. The spectra were recorded at 500 MHz and at 30 °C; the protein concentration was ~8 mM and the sample pH was 3.0-3.2 (uncorrected for deuterium isotope effects).

residue, 83 were between resonances of residues adjacent in the sequence⁹ and 57 were between resonances of non-adjacent residues. NOEs in this last class are shown in Fig. 1. In addition, distance restraints were included corresponding to the 9 hydrogen bonds that could be inferred from observation of slowly exchanging hydrogen atoms and consideration of the secondary structure. Restraints were also included for the 3 disulphide bridges. The 23 phi angles for which ranges could be derived unambiguously (³J_{NH- α CH} > 8 Hz) were included, both as distance and as chirality restraints.

Using distance geometry, five structures consistent with the input constraints and displaying the same 'handedness' were calculated. (When the input distance constraints are fairly imprecise distance geometry techniques can produce structures in which the protein backbone folding is, roughly speaking, the mirror image of the correct fold, although the chirality at each chiral centre is correct. One of these two patterns of chain folding gave consistently lower energy and these results only are given here.) The average of the root-mean-square deviations between the atomic coordinates of pairs of structures is 2.01 $\pm$ 0.18 Å for backbone atoms and 3.42 $\pm$ 0.19 Å for all C, O, N and S atoms. Each structure was subjected to restrained energy minimization and restrained molecular dynamics for a period corresponding to 10-15 ps. The potential energies of these structures at each stage are given in Table 1 together with the restraints energy derived from the NMR distance violations. The molecular

Table 1 The energies of the hEGF structures

Structure	After distance geometry		After energy minimization		After molecular dynamics	
	Potential energy	Restraints energy	Potential energy	Restraints energy	Potential energy	Restraints energy
I	>10 ⁵	63	-1,424	211	-2,222	118
II	>10 ⁵	33	-1,097	214	-1,981	140
III	>10 ⁵	25	-1,128	187	-2,262	79
IV	>10 ⁵	41	-1,386	154	-2,142	52
V	>10 ⁵	36	-1,494	150	-2,146	125

The units are kJ mol⁻¹, but the zero is arbitrary. The potential energy includes contributions from geometric energy terms, van der Waals and electrostatic interactions; the restraints energy arises from violations of the NMR-derived distance restraints. The force constant for the calculation of the restraints energy is 1,000 kJ mol⁻¹ nm⁻².

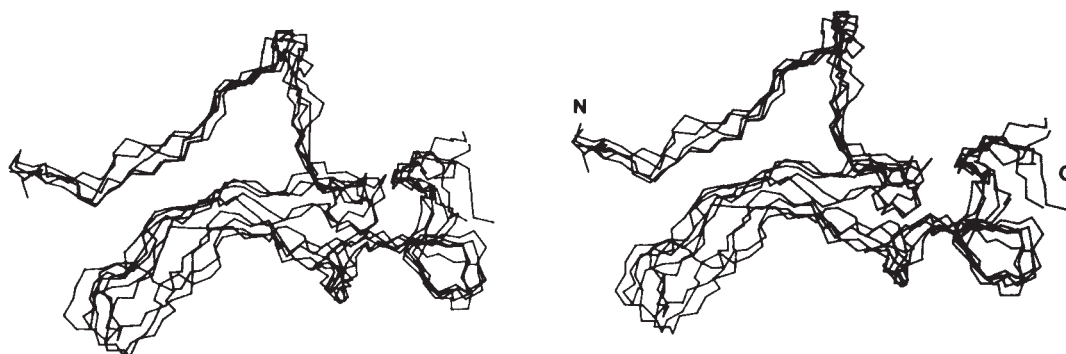


Fig. 2 A stereo view of five distance geometry structures. Only bonds between main chain atoms (N, C α , C') are shown. The structures are overlaid to minimize the differences for residues 3–23 and 28–46.

dynamics simulations had little effect on the overall protein fold, although local conformations were changed. An independent molecular dynamics simulation, starting from an extended structure, produced the same structural features as the distance geometry transformations, including the 'handedness' of the chain fold.

Figure 2 shows stereo views of the structures, overlaid to emphasize the similarities and differences between them. The route followed by the backbone is well reproduced except at the N and C termini and in the turn linking the two strands of the major sheet. Residues 1–32 and 33–48 form two distinct 'domains', which we will refer to as the N- and the C-terminal domains respectively. The first two disulphide bridges (Cys 6–Cys 20 and Cys 14–Cys 31) are in the N-terminal domain whereas the third (Cys 33–Cys 42) is in the C-terminal domain. Between Cys 6 and Cys 14 the chain makes a large loop and relatively few NOEs are observed between resonances of this loop and other parts of the molecule; thus its orientation with respect to the remainder of the molecule may be relatively poorly defined. A kink in the chain is necessary after Cys 14 to enter the first strand of the β -sheet near residue 19. The two strands of the sheet are linked by a turn involving residues 24–27. The type II β -turn (residues 34–37) is followed by another turn in the region of residue 40 to allow formation of the third disulphide bridge and the minor β -sheet (residues 37–38 and 45–44).

Tyrosines 13, 22 and 29 are in close proximity to each other. Aromatic residues are highly conserved at these positions in EGFs and, except for residue 29, in TGF- α s. Aromatic clusters were suggested from NMR studies of mouse and rat EGF¹⁵.

The N terminus may exist for only a small proportion of its time in the triple-stranded sheet arrangement. This would account for the absence of observable hydrogen bonds and the fact that the observed resonance shifts of amino acids in this region are near the 'random coil' values. The resonances from the N terminus are also relatively narrow, implying a high degree of mobility. If interconverting conformers do exist, caution must, of course, be exercised in describing structure in this region of the molecule. Three forms of rat EGF, with up to three of the N-terminal residues missing, have identical activity to the native protein³. The participation of the N terminus in a β -sheet is thus unlikely to be significant for the activity of EGFs.

Spectra of native (1–53) hEGF have also been collected and the same secondary structural features for residues 1–48 are observed. We can therefore predict that the structure of residues 1–48 in the native hEGF is very similar to that described here. Residues 49–53 appear to contribute little to the biological effects of EGFs: rat EGFs, in which these residues are absent, are reported to have full biological activity³ and studies with the 1–47 derivative of hEGF showed it to be equipotent with the native protein *in vivo*¹⁶.

Peptides encompassing residues 20–31 of mouse EGF were reported to bind to the EGF receptor¹⁷ and these residues were also believed to constitute an antigenic region¹⁷. In the structure reported here these residues form part of the major β -sheet, which may thus be a functionally important feature of hEGF. Loss of stability¹⁸ and activity⁷ upon cleaving the peptide bond between residues 21 and 22 is explainable as disruption of this major β -sheet.

EGF-like sequences have been noted in many functionally diverse proteins^{8,19,20} (Fig. 3). For example, EGF-like units in several proteins of the blood clotting system contain the β -hydroxyaspartate residue thought to participate in calcium binding^{21,22}. Site-directed mutagenesis studies have shown the importance of this residue and two aspartate residues for the activity of Factor IX²³. If such units adopt the triple-stranded β -sheet conformation (Fig. 3b) these three acidic residues must lie in close proximity on one face of the sheet, where they could easily form a calcium binding site; calcium may stabilize the triple-stranded sheet. Further, the Tyr/Phe residue thought to be necessary for aspartate hydroxylation²⁴ lies opposite the modified residue (Fig. 3b) and on the same face of the sheet; this may help explain the requirement of the aromatic residue for hydroxylation.

This is a project of the Oxford Enzyme Group which is supported by the SERC. G. D. Searle supplied the crude hEGF and provided financial support for R.M.C.; A.P. received financial support from the MRC. The contributions of J. A. Carver, G. Esposito and S. Wiltshire to this project were invaluable. We

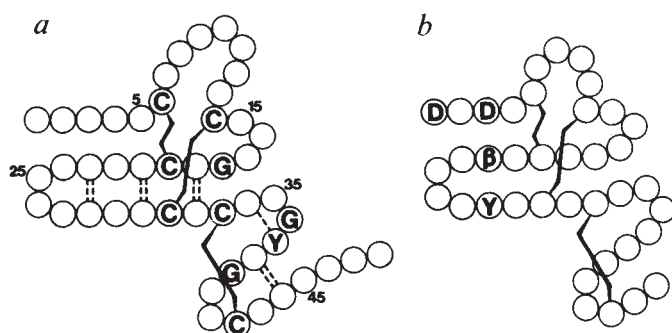


Fig. 3 *a*, Schematic representation of the secondary structural features of (1–48) hEGF. Disulphide bridges are marked with solid lines and hydrogen bonds with dashed lines. EGF-like sequences are found in several proteins and some highly conserved residues are marked; the residue at position 37 is sometimes Phe in the related sequences. The conservation of residues argues strongly for the conservation of the associated structural features. *b*, Extrapolation of the features in *a* to the first EGF-like region of Factor IX (residues 47–84) and Factor X (residues 46–83). Several residues, in addition to those shown in *a*, are highly conserved in Factor IX^{25,26}, Factor X²¹, Protein C^{21,27}, Protein S^{24,28} and complement protein C1r²⁹ and some of those are marked. 'B' represents β -hydroxyaspartate or β -hydroxyasparagine and Phe may replace the marked Tyr.

also thank J. Boyd and N. Soffe for their assistance with the NMR experiments and the Oxford Molecular Biophysics department for the use of computing facilities. The DISGEO program was obtained through the QCPE and the GROMOS program was a gift from W. F. van Gunsteren.

Received 26 January; accepted 7 April 1987.

1. Gregory, H. *Nature* **257**, 325-327 (1975).
2. Cohen, S. *J. biol. Chem.* **237**, 1555-1562 (1962).
3. Simpson, R. J. *et al. Eur. J. Biochem.* **153**, 629-637 (1985).
4. Schlessinger, J. *et al. CRC Crit. Rev. Biochem.* **14**, 93-111 (1983).
5. Marquardt, H. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 4684-4688 (1983).
6. Marquardt, H. & Todaro, G. J. *J. biol. Chem.* **257**, 5220-5225 (1982).
7. Anzano, M. A. *et al. Cancer Res.* **42**, 4776-4778 (1982).
8. Doolittle, R. F. *Trends biochem. Sci.* **10**, 233-237 (1985).
9. Carver, J. A. *et al. FEBS Letts* **205**, 77-81 (1986).
10. Montelione, G. T., Wüthrich, K., Nice, E. C., Burgess, A. W. & Scheraga, H. A. *Proc. natn. Acad. Sci. U.S.A.* **83**, 8594-8598 (1986).
11. Williamson, M. P., Havel, T. F. & Wüthrich, K. *J. molec. Biol.* **182**, 295-315 (1985).
12. Kaptein, R., Zuiderweg, E. R. P., Scheek, R. M., Boelens, R. & van Gunsteren, W. F. *J. molec. Biol.* **182**, 179-182 (1985).
13. Crippen, G. M. & Havel, T. *Acta Crystallogr. sect. A* **34**, 282-284 (1978).
14. van Gunsteren, W. F. & Berendsen, H. J. C. *Biochem. Soc. Trans.* **10**, 301-305 (1982).
15. Mayo, K. H., Schaudies, P., Savage, C. R., De Marco, A. & Kaptein, R. *Biochem. J.* **239**, 13-18 (1986).
16. Hollenberg, M. D. & Gregory, H. *Molec. Pharmacol.* **17**, 314-320 (1980).
17. Komoriya, A. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 1351-1355 (1984).
18. Holladay, L. A., Savage, C. R., Cohen, S. & Puett, D. *Biochemistry* **15**, 2624-2633 (1976).
19. Dayhoff, M. O. *Atlas of Protein Sequence and Structure* Vol. 5, suppl. 3, p. 85 (Natl. Biomed. Res. Found., Washington, 1978).
20. Günzler, W. A. *et al. Hoppe-Seyler's Z. physiol. Chem.* **363**, 1155-1165 (1982).
21. McMullen, B. A. *et al. Biochemistry* **22**, 2875-2884 (1983).
22. Esmon, N. L., DeBault, L. E. & Esmon, C. T. *J. biol. Chem.* **258**, 5548-5553 (1983).
23. Rees, D. J. G. D. *Phil. thesis*, University of Oxford (1986).
24. Stenflo, J., Lundwall, Å. & Dahlbäck, B. *Proc. natn. Acad. Sci. U.S.A.* **84**, 368-372 (1987).
25. Kurachi, K. & Davie, E. W. *Proc. natn. Acad. Sci. U.S.A.* **79**, 6461-6464 (1982).
26. McMullen, B. A., Fujikawa, K. & Kiesel, W. *Biochem. biophys. Res. Comm.* **115**, 8-14 (1983).
27. Fernlund, P. & Stenflo, J. *J. biol. Chem.* **257**, 12170-12179 (1982).
28. Lundwall, Å. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 6716-6720 (1986).
29. Arlaud, G. J., Willis, A. C. & Gagnon, J. *Biochem. J.* **241**, 711-720 (1987).

Reconstitution of a phospholipid flippase from rat liver microsomes

Jonathan M. Backer* & Eliezar A. Dawidowicz†‡

* Division of Health Sciences and Technology, Harvard Medical School, Boston, Massachusetts 02115, USA

† Department of Physiology, Tufts Medical School, Boston, Massachusetts 02111, USA

The endoplasmic reticulum is the principal site of synthesis and initial incorporation of membrane lipids in eukaryotic cells¹; the enzymes of glycerolipid biosynthesis are exclusively located on its cytoplasmic surface². To maintain a phospholipid bilayer in this organelle, newly synthesized phospholipids must be translocated to the luminal surface. Consistent with this are measurements indicating that movement of phospholipids across microsomal membranes is rapid, with a half-time less than 5 min (refs 3 and 4). Rapid movement of phospholipids has also been detected across the plasma membrane of *Bacillus megaterium*^{5,6}, another site of *de novo* lipid biosynthesis. The rapid transmembrane movement of phosphatidylcholine has not been detected, however, in vesicles prepared from microsomal lipids⁴. These latter data suggest involvement in the endoplasmic reticulum of a phospholipid-translocating protein, as was first proposed by Bretscher⁷ who called it 'flippase'. Here we report reconstitution of a phospholipid flippase from rat liver microsomes into lipid vesicles.

The rapid transmembrane movement of phospholipids described above is in marked contrast to their relatively slow

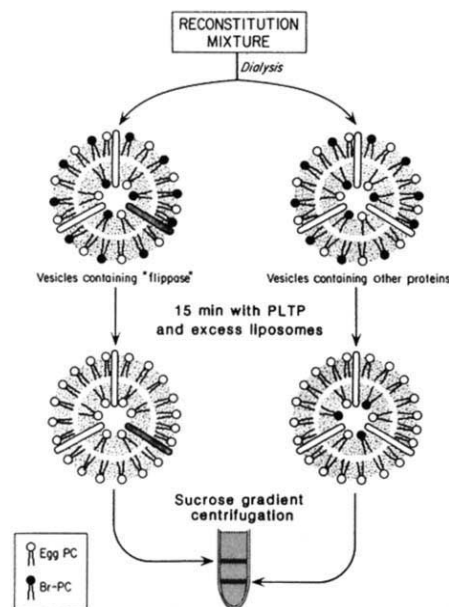


Fig. 1 Outline of reconstitution experiments. Specific materials used are described in 'methods' below. Reconstituted vesicles were prepared by solubilizing 0.5 mg protein (derived from either rat liver microsomes or human erythrocyte ghosts) and a lyophilized mixture of 0.5 mg egg PC, 1.25 mg Br-PC, 0.6 mg egg phosphatidyl-ethanolamine (PE) and 10 μ Ci of ³H-cholesteryl hexadecyl ether (a nonexchangeable vesicle marker¹⁸), in the presence of 3% cholate in dialysis buffer (20 mM NaCl, 5 mM Tris-HCl, 0.5 mM EDTA and 5 mM β -mercaptoethanol pH 6.5) to a final volume of 250 μ l. A lipid control was prepared without added protein. The ratios of both lipid:protein and lipid:detergent¹⁹ used were found to be optimal for reconstitution of the flippase. After standing on ice for 30 min the mixture was centrifuged at 100,000g for 30 min at 4 °C. The resultant supernatant was dialysed for 4 days with four changes of 2 l dialysis buffer at 4 °C using Spectrapor 2 dialysis tubing. The dialysate was layered over 3 ml of 10% sucrose on a 1-ml 60% sucrose cushion and centrifuged at 49,000 r.p.m. for 2 h at 4 °C in an SW 50.1 rotor. The gradient was collected from below and peak fractions at the 10-60% sucrose interface were pooled. An aliquot of these reconstituted vesicles (20-120 nmol total PC in ~15-90 μ l) was incubated for 15 min at 37 °C with partially purified PC-specific PLTP from bovine liver (1 ml) and sonicated liposomes (2-12 μ mol lipid phosphorous in 60-100 μ l) composed of egg PC and bis-dibromostearoyl PE (Br-PE) (1:1 by weight in dialysis buffer). The amount of lipid used in an incubation was dependent on the activity of each PLTP preparation. Typically, preparations had an exchange activity of 100-300 nmol PC min⁻¹ ml⁻¹ at 37 °C as measured in a liposome-mitochondria assay²⁰. Based on this, we used a 50-75-fold excess of PLTP (in 1 ml) over that required to exchange all PC in the aliquot of reconstituted vesicles. Sonicated liposomes were in 50-fold excess. The reaction was stopped on ice and analysed as in Fig. 2.

Methods. Egg PC (type XIE, Sigma), egg PE, Br-PC and Br-PE (Avanti) were shown to be greater than 99% pure by HPLC²¹. The [cholesteryl-1,2-³H(N)]-hexadecyl ether (specific activity 47 Ci mmol⁻¹, NEN) was purified by silicic acid chromatography in hexane/chloroform (3:1 v/v) to a radiochemical purity of greater than 99.8%. Cholic acid (Sigma) was recrystallized three times from 75% ethanol. Labelled ³H-cholic acid (specific activity 25 Ci mmol⁻¹, NEN and Tran³⁵S-label (specific activity 1,000 Ci mmol⁻¹, ICN Radiochemicals) were used without further purification. SM2 Biobeads (Biorad) were washed and stored under water as described by Holloway²²; before use beads were washed into dialysis buffer. Microsomes were prepared from rat liver²³ and washed in 10 mM Tris-HCl. Human erythrocyte ghosts were prepared from outdated blood²⁴. After protein determination²⁵, these membranes were stored at -20 °C for no longer than 2 weeks before use. PC-specific PLTP from bovine liver was purified through the DEAE chromatography step (eluted in 15 mM NaCl, 5 mM sodium phosphate, 10 mM β -mercaptoethanol pH 7.2)²⁰, concentrated using Aquacide IIa (Calbiochem) and stored at 4 °C. Sonicated liposomes were prepared as previously described²⁶.

‡ To whom correspondence should be addressed.

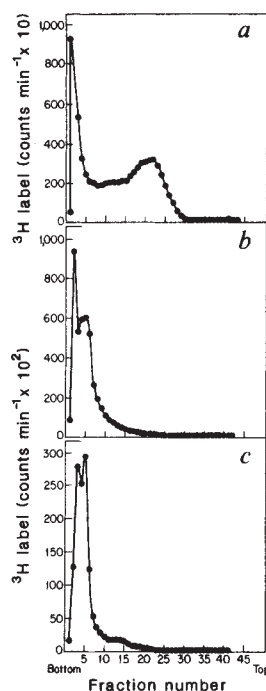


Fig. 2 Typical gradient profiles from incubations performed with reconstituted vesicles containing: *a*, microsomal protein; *b*, human erythrocyte protein; *c*, no added protein.

Methods. Incubation mixtures obtained as described in Fig. 1, (total volume 1.1 ml) were layered onto 10 ml of a 5–20% linear sucrose gradient on a 0.5 ml cushion of 50% sucrose and centrifuged at 39,000 r.p.m. for 24 h at 4 °C in an SW 41 rotor. Under these conditions of centrifugation, the large pool of sonicated egg PC/Br-PE liposomes bands between fractions 1 and 4 (data not shown). Fractions were collected from below and counted using Ultrafluor (National Diagnostics) in a Beckman scintillation counter.

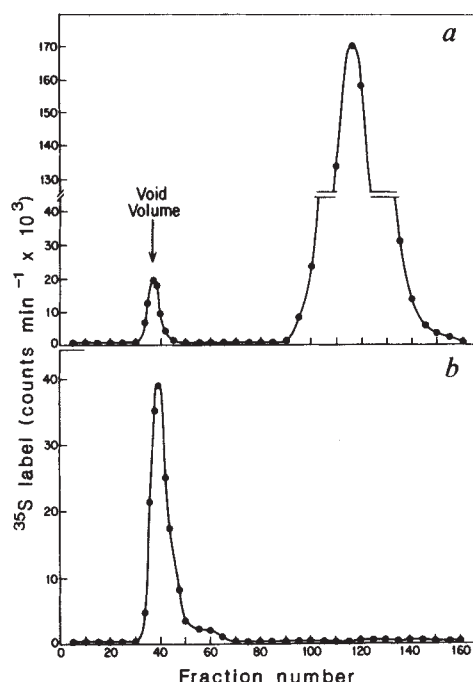


Fig. 3 *a*, Trapping of a marker in reconstituted vesicles containing flippase activity. *b*, Reconstituted vesicles remain sealed during incubation with transfer protein.

Methods. In *a*, an aliquot of reconstituted vesicles (0.58 μ mol PC in 100 μ l), containing microsomal protein and shown to possess flippase activity as in Fig. 2*a*, was permeabilized by the addition of deoxycholate to a final concentration of 0.02% (w/v in dialysis buffer), from a stock solution of 0.05% (w/v), in the presence of 200 μ Ci Tran³⁵S-label²⁷. To remove the deoxycholate this mixture was incubated with 1 ml of a suspension of SM2 Biobeads (50% by volume in dialysis buffer) at 4 °C overnight on a rocking platform²⁸. After removal of the Biobeads by filtration through glass wool, untrapped Tran³⁵S-label was removed by chromatography on Sephadex G-50 (2.5 $\times$ 20 cm). Fractions (0.5 ml) were collected and 25 μ l aliquots counted as in Fig. 2. In *b*, fractions eluting at the void volume (from *a*) were pooled. The vesicles obtained (14 nmol PC in 400 μ l) were incubated with 1 ml of the PC-specific PLTP and egg PC/Br-PE sonicated liposomes for 15 min at 37 °C as described in Fig. 1. The resultant mixture was chromatographed on Sephadex G-50 as before. No leakage of trapped marker was observed, in marked contrast to the leakage observed after incubation with 0.05% (w/v) deoxycholate for 15 min at 37 °C (data not shown).

movement across other membranes (with half-times of 2–4 h, ref. 8). Attempts to induce rapid movement of phospholipids across synthetic vesicles by incorporation of integral membrane proteins derived from the human erythrocyte^{9,10}, from asymmetrical perturbation of the bilayer by digestion with phospholipase D¹¹, or by oxidation of the phospholipids¹² have achieved half-times of the order of one hour. However, the reconstitution of rapid transmembrane movement of phosphatidylcholine (PC), characteristic of microsomal membranes, has not been previously reported.

Our approach (see Fig. 1) is influenced by the work of Goldin and Rhoden¹³, who used the activity of transport proteins to alter the physical characteristics of vesicles into which these proteins had been reconstituted, thereby facilitating their separation and partial purification. We have reconstituted total microsomal protein into lipid vesicles containing a density-labelled

phospholipid, bis-dibromostearoyl PC (Br-PC)¹⁴. Incubation of these vesicles with a PC-specific transfer protein and liposomes containing 'light' egg PC was used to alter selectively the density of vesicles containing a flippase. Phospholipid transfer protein (PLTP), which cannot pass through membranes, will equilibrate PC in the outer leaflet of different membranes^{15,16}; equilibration of PC in the inner leaflet depends on the rate of transmembrane movement. A vesicle containing a flippase can therefore be identified by the complete availability of its lipids for exchange. When 'dense' vesicles containing a flippase are incubated with PLTP and liposomes containing light PC, all the dense PC should be exchanged for light. In vesicles without a flippase, only the dense PC in the outer leaflet will be exchanged for light and these vesicles will remain relatively dense. The two populations can then be separated by density gradient centrifugation.

Incubation of vesicles reconstituted with microsomal protein as described (see Fig. 1 legend) resulted in a shift of 52% of these vesicles to a lower density (Fig. 2a). We have obtained similar results using ten separate preparations. Omission of PLTP resulted in a single vesicle peak at the bottom of the gradient (data not shown). In experiments using vesicles reconstituted with human erythrocyte ghosts or with no added protein, comparable shifts in density were not seen (Fig. 2b and c). These data demonstrate reconstitution of a flippase activity from rat liver microsomes.

In control experiments ^3H -cholate was included in the initial solubilization. After dialysis, residual levels were less than 0.001% cholate by weight. We have demonstrated that cholate at this concentration has no effect on the rate of transbilayer movement of PC (data not shown). Furthermore, the density shift in microsomal preparations (Fig. 2a) is not a result of vesicle leakiness, allowing direct access of PLTP to inner leaflet lipids, because a trapped marker is retained during incubations with PLTP (Fig. 3b).

Controls were conducted to show that the position of reconstituted microsomal vesicles on sucrose gradients after density shift corresponded to that of complete exchange of dense lipid for light. To achieve this, 0.02% cholate was included during an exchange incubation; at this detergent concentration, all the lipid in sonicated liposomes becomes available for exchange (data not shown). The position of the density-shifted vesicles after this control incubation corresponds closely to that observed in Fig. 2a (data not shown).

This is the first study to demonstrate the reconstitution into lipid vesicles of a PC translocase activity whose velocity is comparable to that seen in rat liver microsomes. We have achieved complete exchange of PC in 15 min from a population of vesicles containing microsomal protein. Flippase activity was not detected in reconstituted vesicles containing extracted microsomal lipid (data not shown) suggesting that the translocase activity is protein mediated. Because we have not detected comparable activity using proteins from human erythrocyte ghosts, this would suggest that rapid transbilayer movement of PC in microsomal membranes is due to the presence of a specific microsomal protein. We estimate that our results could be explained by the presence of flippase among total microsomal protein at a minimum level of 0.6% by weight. However, we cannot exclude the possibility that flippase is present at higher concentrations.

In a related study¹⁷ it has been suggested that transport of short-chain water soluble PCs across microsomal membranes is a specific protein-mediated process which is sensitive to treatment with trypsin and *N*-ethylmaleimide. By contrast, we have demonstrated reconstitution of a translocase activity which acts upon long-chain bilayer forming PCs¹⁴ and is insensitive to these reagents under identical conditions (data not shown). Our data are evidence for a phospholipid-translocating protein in rat liver microsomes which could be involved in the biogenesis of the phospholipid bilayer of the endoplasmic reticulum. In addition, the methods described suggest a possible scheme for the purification of this protein.

This work was supported by grants from NIH, the American Heart Association and the Charlton Fund of Tufts Medical School. We thank Dr Stanley Goldin for helpful discussions.

Received 9 March; accepted 2 April 1987.

10. Gerritsen, W. J., Henricks, P. A. J., de Kruijff, B. & Van Deenen, L. L. M. *Biochim. biophys. Acta* **600**, 607-619 (1980).
11. De Kruijff, B. & Baken, P. *Biochim. biophys. Acta* **507**, 38-47 (1978).
12. Shaw, J. M. & Thompson, T. E. *Biochemistry* **21**, 920-927 (1982).
13. Goldin, S. M. & Rhoden, V. J. *biol. Chem.* **253**, 2575-2583 (1978).
14. Dawidowicz, E. A. & Rothman, J. E. *Biochim. biophys. Acta* **455**, 621-630 (1976).
15. Johnson, L. W., Hughes, M. E. & Zilversmit, D. B. *Biochim. biophys. Acta* **375**, 176-185 (1975).
16. Rothman, J. E. & Dawidowicz, E. A. *Biochemistry* **14**, 2809-2816 (1975).
17. Bishop, W. R. & Bell, R. M. *Cell* **42**, 51-60 (1985).
18. Pool, G. L., French, M. E., Edwards, R. A., Huang, L. & Lumb, R. H. *Lipids* **17**, 448-452 (1982).
19. Klausner, R. D., Van Renswoude, J., Blumenthal, R. & Rivnay, B. in *Molecular and Chemical Characterization of Membrane Receptors* (eds Venter, J. C. & Harrison, L. C.) 209-239 (Liss, New York, 1984).
20. Westerman, J., Kamp, H. H. & Wirtz, K. W. A. *Meth. Enzym.* **98**, 581-586 (1983).
21. Chen, S. S.-H. & Kou, A. Y. J. *Chromatogr.* **227**, 25-31 (1982).
22. Holloway, P. W. *Analyt. Biochem.* **53**, 304-308 (1973).
23. Eriksson, L. C. *Biochim. biophys. Acta* **508**, 155-164 (1978).
24. Hanahan, D. J. & Ekholm, J. E. *Meth. Enzym.* **31**, 168-172 (1974).
25. Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. *J. biol. Chem.* **193**, 265-275 (1951).
26. Backer, J. M. & Dawidowicz, E. A. *Biochim. biophys. Acta* **551**, 260-270 (1979).
27. Goldin, S. M., Rhoden, V. & Hess, E. J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6884-6888 (1980).
28. Lorusso, D. J. & Green, F. A. *Science* **188**, 66-67 (1975).

Catalysis of splicing-related reactions between dinucleotides by a ribozyme

Peter S. Kay & Tan Inoue

The Salk Institute for Biological Studies, PO Box 85800, San Diego, California 92138, USA

Certain intervening sequence (IVS) RNAs catalyse their own excision from a primary transcript to yield mature RNA in a reaction termed self-splicing¹⁻⁴. These Group I IVS RNAs contain several conserved sequences and possess a common secondary structure⁵⁻⁷. The *Tetrahymena* precursor ribosomal RNA possesses an IVS of this group that is known to self-splice *in vitro*. The nature of this IVS is of great interest to both biology and chemistry, because understanding its catalytic activity should shed new light on the function of RNA in biological systems and the evolution of RNA which might be relevant to the early stages of life on Earth^{8,9}. We have analysed the minimum requirement for this reaction as one approach to understanding the mechanism of this RNA catalysis. We now show that a fragment of the IVS RNA of *Tetrahymena* can mediate a simple transesterification reaction between the substrate GpN (where N is A, C, G or U) and the nucleophile CpU. This newly discovered reaction and its reverse reaction represent the fundamental catalytic activity of the self-splicing Group I IVSs.

The IVS of the *Tetrahymena* precursor ribosomal RNA (pre-rRNA) is precisely excised *in vitro* through nucleophilic attack by a monomeric guanosine derivative in the presence of magnesium chloride. This self-splicing reaction proceeds through two separate transesterification reactions¹. In the first step, a monomeric guanosine derivative attacks the 5' splice site to produce the 5' exon and an RNA consisting of the monomeric guanosine derivative, the IVS and the 3' exon. In the second step, the released 5' exon attacks the 3' splice site to complete ligation of the exons. The reactions are mediated by the IVS RNA itself (Fig. 1a).

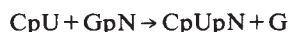
A linear form of the IVS RNA, lacking the 5' monomeric guanosine derivative and the next 18 nucleotides (L-19 IVS RNA), acts as a catalyst which is capable of directing nucleotydyl transfer¹⁰, phosphotransfer¹¹ and a sequence-specific cleavage reaction¹², using oligonucleotides as substrates. All these reactions are initiated by nucleophilic attack by a guanosine residue, either as the nucleotide located at the 3' end of the IVS RNA (G at position 414) or as a monomeric guanosine derivative. These catalytic reactions and the first step of self-

1. Wilgram, G. F. & Kennedy, E. P. *J. biol. Chem.* **238**, 2615-2619 (1963).
2. Bell, R. M., Ballas, L. M. & Coleman, R. A. *J. Lipid Res.* **22**, 391-403 (1981).
3. Zilversmit, D. B. & Hughes, M. E. *Biochim. biophys. Acta* **469**, 99-110 (1977).
4. Van den Besselaar, A. H. M. P., de Kruijff, B., Van den Bosch, H. & Van Deenen, L. L. M. *Biochim. biophys. Acta* **510**, 242-255 (1978).
5. Rothman, J. E. & Kennedy, E. P. *Proc. natn. Acad. Sci. U.S.A.* **74**, 1821-1825 (1977).
6. Langley, K. E. & Kennedy, E. P. *Proc. natn. Acad. Sci. U.S.A.* **76**, 6245-6249 (1979).
7. Bretscher, M. S. in *The Cell Surface in Development* (ed. Moscona, A. A.) 17-27 (Wiley, New York, 1973).
8. Dawidowicz, E. A. *Curr. Top. Membranes Transp.* **29**, 175-202 (1987).
9. De Kruijff, B., Van Zoelen, E. J. J. & Van Deenen, L. L. M. *Biochim. biophys. Acta* **509**, 537-542 (1978).

splicing share a common mechanism that involves transesterification by guanosine. It has been suggested that there is a particular binding site in the IVS which recognizes guanosine^{13,14}.

Because our interest lies in identifying the minimum substrate and catalyst requirements for the self-splicing reaction, as an initial study we attempted a simplified intermolecular version of self-splicing. In this reaction dinucleotides were incubated with cofactor nucleotides in the presence of a linear form of the IVS RNA. As described above, the 3' terminal G 414 of L-19 IVS RNA serves as a nucleophile and might therefore interfere with the interaction of dinucleotides or cofactor nucleotides containing guanosine to the guanosine-specific binding site¹⁵. To avoid the interference by G 414, a large fragment of the IVS RNA lacking G 414 was used in the reactions. This fragment, termed L-21 *ScaI* IVS RNA, lacks the five nucleotides at the 3' end of the IVS as well as the 5' monomeric guanosine derivative and the next 20 nucleotides (Fig. 1b). It is known that GTP can act as a nucleophile in the first step of self-splicing and that CpU (the last two nucleotides of the 5' exon sequence) can act as a 5'-exon-equivalent by intermolecularly ligating with the 3' exon¹⁶. Thus each of the sixteen 5' end-[³²P]-labelled dinucleotides was incubated with GTP, CpU or no cofactor in the presence of L-21 *ScaI* IVS RNA. The results are shown in Fig. 2.

All the dinucleotides [³²P]pGpN (where N is A, C, G or U) (unless otherwise stated the 5' phosphate of oligonucleotides is labelled) were cleaved in the presence of CpU to produce [³²P]pG (Fig. 2c). No such cleavage was detected when the reaction was attempted with the other twelve dinucleotides (Fig. 2a, b and d). The reaction was also attempted using [³²P]pCpU and non-labelled GpA, GpC, GpG or GpU in the presence of L-21 *ScaI* IVS RNA. The four GpN dinucleotides were readily converted to labelled trinucleotides (Fig. 3a), which we believe to be of the form [³²P]pCpUpN on the basis of sequence analysis (data not shown). These results confirm that L-21 *ScaI* IVS RNA is capable of mediating the following transesterification reaction between CpU and GpN:



The 3' splice site preceded by G 414 in the precursor ribosomal RNA is cleaved by CpU in exactly the same manner in a reaction analogous to exon ligation. Therefore, the reaction between CpU and GpN represents a mini-exon ligation reaction in which CpU acts as the 5' exon and the phosphodiester bond of GpN acts as the 3' splice site.

The labelled dinucleotide [³²P]pGpG was specifically hydrolysed to produce [³²P]pG at pH 7.5 (Fig. 2c). Although the efficiency was much lower, the specific hydrolysis of [³²P]pGpA, [³²P]pGpC and [³²P]pGpU were also detected at this pH (data not shown). It is known that the analogous cleavage reaction occurs at the 3' splice site in the precursor ribosomal RNA, yielding a 2', 3'-OH and a 5' phosphate at the newly created termini¹⁴. This reaction has been shown to take place more quickly at higher pH. When the specific hydrolysis of the four [³²P]pGpNs was attempted in the presence of L-21 *ScaI* IVS RNA at pH 8.5, an increase in hydrolysis with increased pH was clearly detected for all four dinucleotides (data not shown). We conclude that all four pGpNs can be hydrolysed as a 3'-splice-site equivalent when incubated in the presence of L-21 *ScaI* IVS RNA.

Combining the results for the reactions of pGpN, we see that the phosphodiester bond preceded by guanosine in the dinucleotides possesses two unique reactivities which are specific to the 3' splice site. These results provide evidence that a guanosine preceding a phosphodiester bond is the minimum requirement for the IVS RNA to specify a 3'-splice-site equivalent. This guanosine is presumably recognized by a guanosine-specific binding site as in the first step of splicing^{13,17,18}.

However, in self-splicing of the *Tetrahymena* pre-rRNA only the phosphodiester bond preceded by G 414 is distinguished as

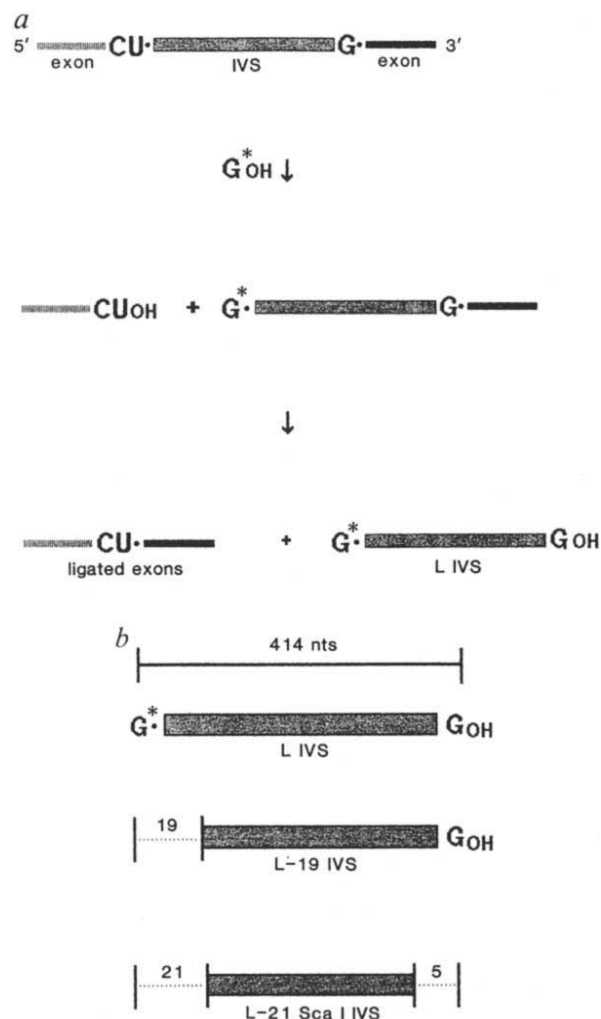


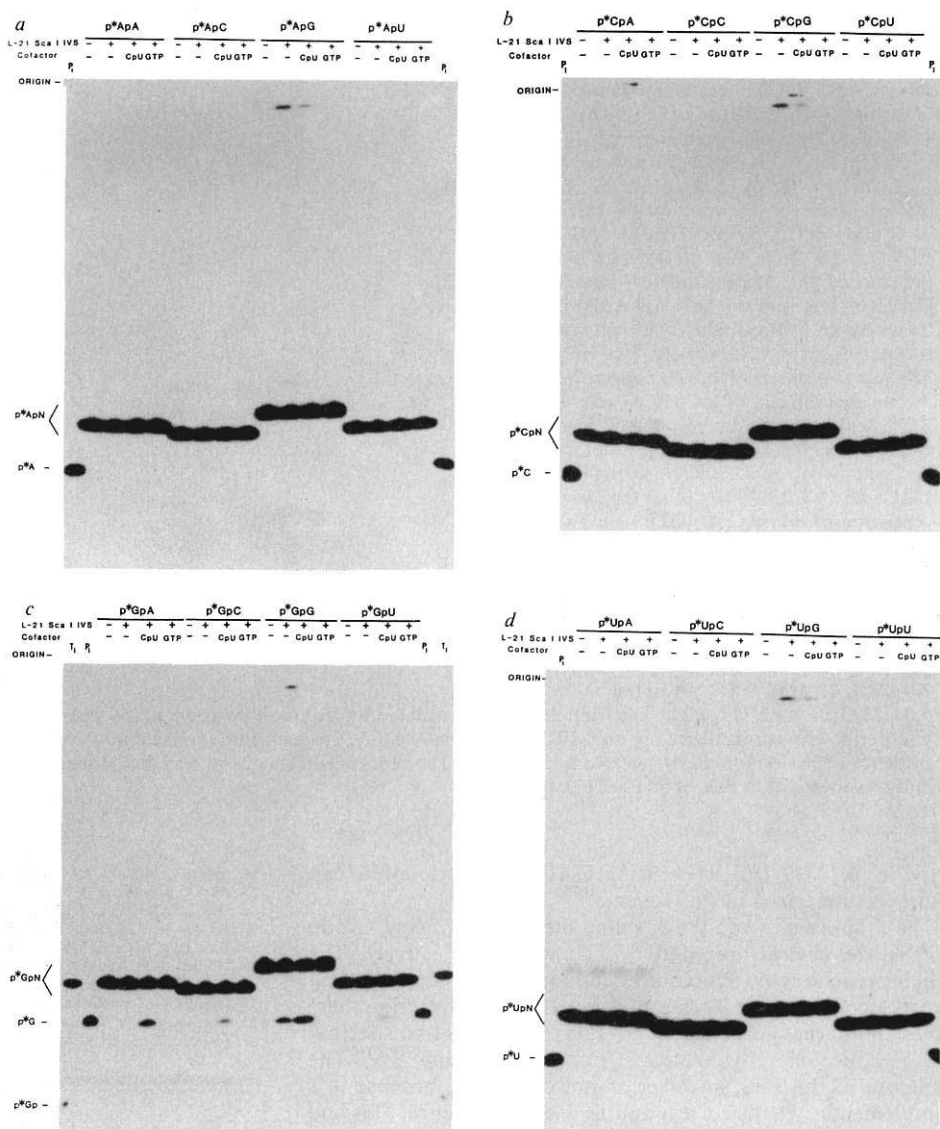
Fig. 1 The *Tetrahymena* self-splicing reaction and linear forms of the IVS RNA. *a*, Schematic model of the two-step self-splicing reaction. G*OH represents the monomeric guanosine derivative that attacks the 5' splice site in the first step. *b*, Linear forms of the IVS. L IVS represents the excised IVS RNA (414 nucleotides (nt)) that is the direct product of precursor ribosomal RNA splicing. L-19 and L-21 *ScaI* IVS represent shortened versions of the L IVS RNA. The sequence at the termini of L-21 *ScaI* IVS RNA was determined by enzymatic analysis of the 5' and 3' end-labelled RNAs^{22,23}. (The results of this analysis identified the 5' end nucleotide as G 22 and the 3' end nucleotide as U 409).

the 3' splice site from among several potential sites. What specifies this phosphodiester bond as the 3' splice site and prevents other potential sites preceded by guanosine from being selected? This question indicates that 3' splice site selection is governed by an additional mechanism in conjunction with the recognition of guanosine¹⁶. We believe that this mechanism involves either the recognition of the sequence near the 3' splice site or the location of G 414 relative to its recognition site in the tertiary structure of the pre-rRNA.

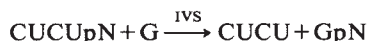
In the presence of L-19 IVS RNA, RNA substrates containing the sequence CUCU were tested for the presence of 5'-splice-site equivalents in the presence of a monomeric guanosine derivative¹². It was found that all the substrates were specifically cleaved by GTP after the sequence CUCU, indicating that the CUCU sequence is necessary and sufficient to designate a phosphodiester bond as the 5' splice site in the *Tetrahymena* pre-rRNA. The reaction is equivalent to the first step of splicing

Fig. 2 Reactions of L-21 *ScaI* IVS RNA with 16 5'-end-labelled dinucleotides. Each 5'-end-labelled dinucleotide (0.5 μ M) was incubated either alone as a control or with L-21 *ScaI* IVS RNA (0.6 μ M) in the presence of CpU (5 mM), GTP (5 mM), or no cofactor. The reaction was carried out in 50 mM $MgCl_2$, 50 mM HEPPS (pH 7.5) at 42 °C for 2 h. *a*, [32 P]pApN; *b*, [32 P]pCpN; *c*, [32 P]pGpN; *d*, [32 P]pUpN, (N = A, C, G or U). Note that the asterisk indicates the labelled phosphate. T_1 denotes treatment of [32 P]pGpA with ribonuclease T_1 . P_1 denotes treatment of [32 P]pNpA with ribonuclease P_1 . The bands appearing near the origin for the reactions of [32 P]pNpG with L-21 *ScaI* IVS presumably represent [32 P]pNpG addition products at or near the 5' end of L-21 *ScaI* IVS RNA, although the sequences of these products have not yet been identified. Note that the $MgCl_2$ concentration of 50 mM represents optimum reaction conditions. The reactions described here and in Fig. 3 were also seen to proceed at 5 mM $MgCl_2$, although less efficiently.

Methods. A linear form of the *Tetrahymena* precursor ribosomal RNA was synthesized by transcription *in vitro* of the plasmid pT7TT1A3 which had been truncated with *ScaI* endonuclease¹². L-21 *ScaI* IVS RNA was prepared by incubating the transcript with CpU (5 mM) and GTP (5 mM) in 50 mM $MgCl_2$, 50 mM HEPPS (pH 7.5) at 42 °C for 2 h. The sixteen dinucleotides (Sigma) were 5'-end-labelled by incubating 1 mM of the dinucleotide, 1 μ M of [γ - 32 P]ATP (4,000–5,000 Ci mmol⁻¹) and 2 units μ l⁻¹ T4 polynucleotide kinase in 5 mM $MgCl_2$, 3 mM DTT, 25 mM CHES (pH 9.5) at 37 °C for 90 min. Each labelled dinucleotide ran as a single band on a 20% polyacrylamide gel, except for trace impurities as seen with [32 P]pApC, [32 P]pCpA, and [32 P]pUpA.

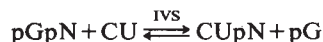


and can be schematically represented as follows:



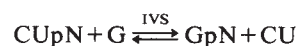
(N represents A, C, G or U; p represents the phosphodiester bond at the reaction site).

In our experiments, however, GTP did not react with any of the sixteen dinucleotides. This demonstrates that a dinucleotide cannot act as a substrate in the reaction equivalent to the first stage of splicing. Yet, because the first and second steps of splicing involve the recognition of guanosine, we postulate that our simplified transesterification reaction may represent not only mini-exon ligation but could also be regarded as the reverse of the first step of self-splicing:



This scheme would indicate that the trinucleotide CUN (containing a phosphodiester bond equivalent to the 5' splice site)

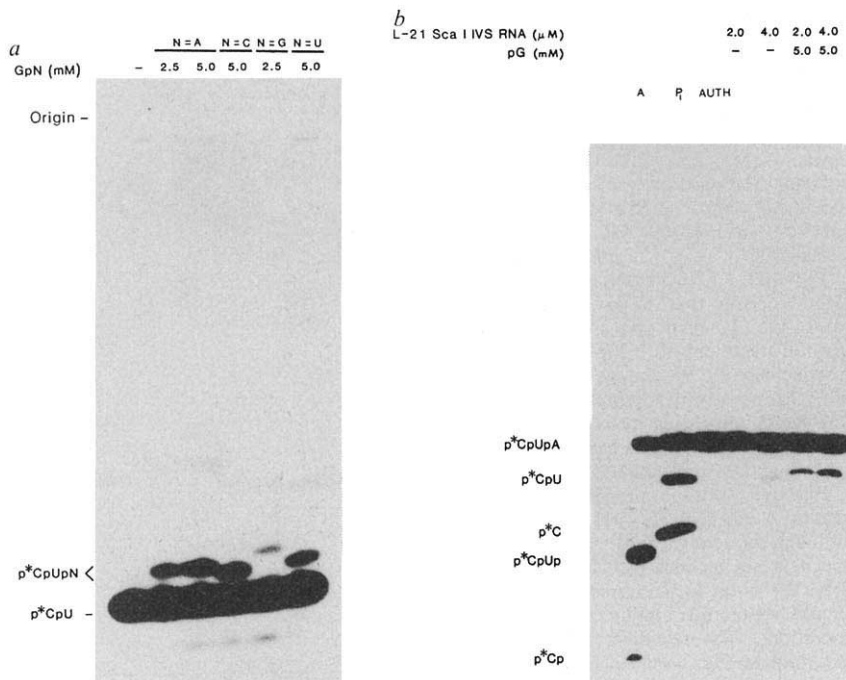
should work as a substrate with a monomeric guanosine derivative. Thus a 5' end-labelled trinucleotide, [32 P]pCpUpA, was isolated and reacted with pG in the presence of L-21 *ScaI* IVS RNA to investigate this reverse reaction (equivalent to the first stage of self-splicing). As shown in Fig. 3b, [32 P]pCpUpA was cleaved in the presence of pG to produce [32 P]pCpU. This result demonstrates that CU preceding a phosphodiester bond is the minimum sequence required by the IVS to specify that phosphodiester bond as equivalent to the 5' splice site. Therefore the two transesterification reactions required for self-splicing of *Tetrahymena* pre-rRNA can be represented as one simple reaction scheme:



The forward reaction of CUpN with G represents the first step of self-splicing (cleavage by a monomeric guanosine derivative), whereas the reverse reaction of GpN with CU represents the second step of self-splicing (ligation of exons). The catalytic

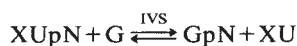
Fig. 3 The mini self-splicing reactions. *a*, Incubation of 5 μ M [32 P]pCpU with either 0.6 μ M L-21 *ScaI* IVS RNA and no additional cofactor (as a control) or with 0.6 μ M L-21 *ScaI* IVS RNA with 2.5 mM and/or 5.0 mM GpN (N = A, C, G, U). The reaction was carried out in 50 mM MgCl₂, 50 mM HEPES (pH 7.5) at 42 °C for 2 h. *b*, Incubation of 0.25 μ M [32 P]pCpUpA with either 2.0 μ M or 4.0 μ M L-21 *ScaI* IVS RNA in the presence of 5 mM pG or no cofactor. The reaction was carried out in 50 mM MgCl₂, 50 mM HEPES (pH 7.5) at 42 °C for 2 h. Note that the asterisk indicates the labelled phosphate. A denotes treatment of [32 P]pCpUpA with ribonuclease A. P_i denotes treatment of [32 P]pCpUpA with ribonuclease P_i. AUTH denotes untreated [32 P]pCpUpA. The minor amount of product formed during the incubation of [32 P]pCpUpA with 4.0 μ M L-21 *ScaI* IVS RNA presumably represents either hydrolysis of the trinucleotide or the 3'-end activity of an IVS RNA copurified with L-21 *ScaI* IVS RNA. Note that the 5 mM concentration of either CpU, GTP or pG used in the experiments described here and in Fig. 2 represents excess nucleophile. The reactions were also seen to proceed at nucleophile concentrations of 50 μ M.

Methods. The labelled [32 P]pCpUpA was prepared by incubating 5 μ M [32 P]pCpU with 2.0 μ M L-21 *ScaI* IVS RNA in the presence of 5.0 mM GpA at 42 °C for 2 h. The trinucleotide was isolated by electrophoresis on a 20% polyacrylamide denaturing gel, eluted from the gel, and purified by chromatography on G10-120. The dinucleotide product from the incubation of [32 P]pCpUpA and L-21 *ScaI* IVS RNA in the presence of pG was similarly isolated and purified. The sequence of this band was determined by enzymatic digestion with RNase P_i and A and was identical to that of [32 P]pCpU.



activities of L-19 IVS RNA using G 414 as a nucleophile can also be summarized in this scheme¹⁰⁻¹².

The 5' splice site and the 3' splice site of self-splicing Group I IVSs are always preceded by U and G, respectively. The splicing reactions always require a monomeric guanosine derivative as a cofactor. In all cases of this self-splicing reaction there exists an internal guide sequence (IGS) in the IVS that specifies the 5' splice site by Watson-Crick base pairing^{1,19,20}. The sequence of the IGS, however, is not conserved, because it is complementary to the corresponding 5' exon sequence. This and data from previous studies²¹ indicate that a dinucleotide consisting of the last two nucleotides of the 5' exon sequence for any self-splicing Group I IVS, XU (X = A, C, G or U), can be recognized by its IVS as a 5'-exon-equivalent. It therefore seems that the self-splicing reaction for all Group I IVSs can be generalized as follows:



We thank L. E. Orgel, G. F. Joyce and T. R. Cech for comments on the manuscript and S. Bailey for manuscript preparation. This work was supported by grants from the National Institutes of Health and the Alfred Krupp von Bohlen und Halbach-Stiftung.

Received 3 March; accepted 31 March 1987.

1. Zaug, A. J., Grabowski, P. J. & Cech, T. R. *Nature* **301**, 578-583 (1983).
2. Garriga, G. & Lambowitz, A. M. *Cell* **38**, 631-641 (1984).
3. Horst, G. v. d. & Tabak, H. F. *Cell* **40**, 759-766 (1985).
4. Chu, F. K., Maley, G. F., West, D. K., Belfort, M. & Maley, F. *Cell* **45**, 157-166 (1986).
5. Davies, R. W., Waring, R. B., Ray, J. A., Brown, T. A. & Scazzocchio, C. *Nature* **300**, 719-724 (1982).
6. Waring, R. B. & Davies, R. W. *Gene* **28**, 277-291 (1984).
7. Michel, F. & Dujon, B. *EMBO J.* **2**, 33-38 (1983).
8. Cech, T. R. *Sci. Am.* **255**, 64-75 (1986).
9. Orgel, L. E. *J. theor. Biol.* **123**, 127-149 (1986).
10. Zaug, A. J. & Cech, T. R. *Science* **231**, 470-475 (1986).
11. Zaug, A. J. & Cech, T. R. *Biochemistry* **25**, 4478-4482 (1986).
12. Zaug, A. J., Been, M. D. & Cech, T. R. *Nature* **324**, 429-433 (1986).
13. Bass, B. L. & Cech, T. R. *Nature* **308**, 820-826 (1984).

14. Inoue, T., Sullivan, F. X. & Cech, T. R. *J. molec. Biol.* **189**, 143-165 (1986).
15. Kay, P. S. & Inoue, T. *Nucleic Acids Res.* **15**, 1559-1577 (1987).
16. Inoue, T., Sullivan, F. X. & Cech, T. R. *Cell* **43**, 431-437 (1985).
17. Tanner, N. K. & Cech, T. R. *Biochemistry* (in the press).
18. Bass, B. L. & Cech, T. R. *Biochemistry* **25**, 4473-4477 (1986).
19. Waring, R. B., Townner, P., Minter, S. J. & Davies, R. W. *Nature* **321**, 133-139 (1986).
20. Been, M. D. & Cech, T. R. *Cell* **47**, 207-216 (1986).
21. Garriga, G., Lambowitz, A. M., Inoue, T. & Cech, T. R. *Nature* **322**, 86-89 (1986).
22. Donis-Keller, H., Maxam, A. M. & Gilbert, W. *Nucleic Acids Res.* **4**, 2527-2538 (1977).
23. England, T. E. & Uhlenbeck, O. C. *Nature* **275**, 560-561 (1978).

Autocatalytic synthesis of a tetranucleotide analogue

Wojciech S. Zielinski & Leslie E. Orgel

The Salk Institute for Biological Studies, PO Box 85800, San Diego, California 92138, USA

As an approach to the study of the kind of chemical process that might have contributed to the origin of life, attempts have been made to develop purely chemical systems in which oligonucleotides self-replicate. Although preformed oligonucleotides have been shown to facilitate the formation of their complements from activated mononucleotides, only a restricted range of oligomers are efficient templates¹ and it will clearly be difficult to find a pair of complementary oligomers each of which will facilitate the formation of the other². Many of the difficulties facing the development of a self-replicating system could be overcome by using a pair of complementary substrate molecules that condense together more easily than ribonucleotides. It would also be helpful if each substrate molecule contained equal numbers of purine and pyrimidine bases as, otherwise, there is a tendency for purines to be over-represented in the products¹. We have therefore explored the chemistry of 3'-amino-3'-deoxynucleotides³ and their dimers⁴. We

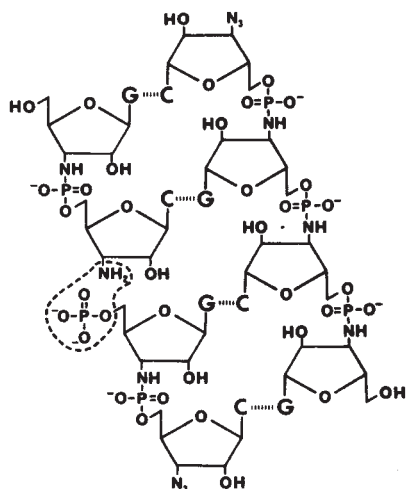


Fig. 1 The mini-helix in which template-directed condensation occurs. Broken bars indicate hydrogen bonds.

report here that the tetranucleoside triphosphoramidate GNhpCNhpGNhpCN₃ acts as a template to catalyse the condensation of GNhpCNH₂ and pGNhpCN₃, forming further molecules of the template. The system is therefore autocatalytic, and in accordance with elementary theory the amount of product made increases with the square root of the template concentration.

The substrate pGNhpCN₃ and ¹⁴C-labelled GNhpCNH₂ (specific activity 2 nCi μmol⁻¹) as well as the tetramer GNhpCNhpGNhpCN₃ (see Fig. 1) were prepared by published procedures⁴. Reactions were carried out in volume of 500 μl at 5°C in polyethylene microcentrifuge tubes. The pH was maintained at 7.5 using a 0.1 M solution of the sodium salt of 4-(2-hydroxyethyl)-1-piperazinesulphonic acid (HEPES) as buffer. Each substrate was always present at a concentration of 0.001 M and the tetramer GNhpCNhpGNhpCN₃ was added at appropriate concentrations. The reactions were initiated by adding 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide (0.06 M final concentration). Aliquots (90 μl) were taken from the reaction mixtures, from time to time, and chromatographed for 23–24 h against authentic markers on Whatmann 3MM paper in *n*-propanol/28% NH₄OH/H₂O (55:10:35). Zones corresponding to the labelled substrate GNhpCNH₂ and the tetramer were cut out, immersed in Liquifluor and counted in a Beckman scintillation counter.

We found that the condensation of ¹⁴C-labelled GNhpCNH₂ with pGNhpCNH₂ under all the conditions used in our experiments gave a single radioactively labelled product. It was shown to be GNhpCNhpGNhpCN₃ by co-chromatography with authentic material. In our experiments on autocatalysis, we chose concentrations of reactants and reaction times so that the amount of product formed in the absence of added tetramer was small, but sufficient to permit accurate scintillation counting. This allowed us to measure autocatalytic synthesis against a small but well-defined background of uncatalysed reaction. The yields of product tetramer formed in a control reaction without template, and in the presence of 0.25 mM, 0.5 mM and 1 mM template are shown in Fig. 2.

The data summarized in Fig. 2 are similar to those reported for the autocatalytic condensation of two suitably protected deoxytrinucleotides, d(5'MeCCGp) and d[CGGp(o-ClPh)], using a water soluble carbodiimide to give the deoxyhexanucleotide d[5'CCGCGGp(o-ClPh)]⁵. The presence of tetramer template substantially increases the rate of production of identical new tetramer molecules. In the reaction mixture containing 0.25 mM tetramer, each template molecule had directed the synthesis of approximately 0.4 new tetramer molecules at the

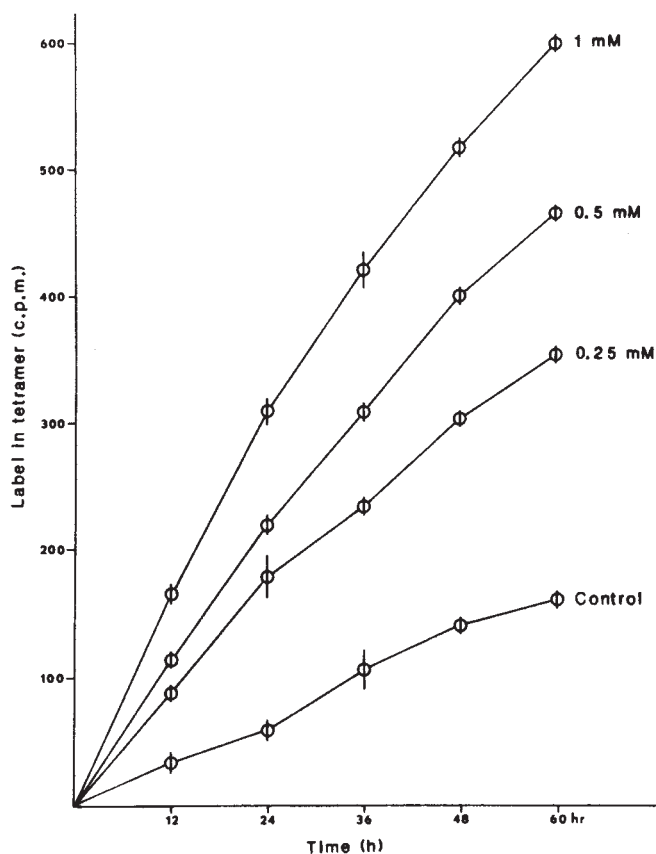


Fig. 2 The dependence of the yield of radioactive tetramer on the initial concentration of non-radioactive tetramer (shown on each curve). Circles, mean of two measurements; the vertices of the bars give the measured values. The concentration of each dimer was always 1 mM. A total of 2,150 c.p.m. was measured in each 90-μl aliquot.

end of the experiment. Within the accuracy of our measurements, the rate of tetramer formation is described by an equation of the form

$$\frac{dC_{\text{tet}}}{dt} = \alpha + \beta\sqrt{C_{\text{tet}}}$$

where C_{tet} is the total concentration of tetramer, and α and β are constants. A roughly square-root dependence of the reaction rate on the total concentration of tetramer follows from the stability of the miniduplex formed by the tetramer⁵.

These results establish for the first time that nucleic-acid-like oligomers with a modified backbone can act as templates for the assembly of complementary oligomers with the same backbone. They also show that there is a concentration range in which the rate of template-directed synthesis using dimers as substrates substantially exceeds the rate of spontaneous oligomerization, even when a tetramer that forms a stable duplex is used as template. These are necessary (but not sufficient) conditions for the development of a non-enzymatic self-replicating system.

We thank Sylvia H. Bailey for manuscript preparation. The work was supported by an NSF grant.

Received 26 January; accepted 31 March 1987.

1. Inoue, T. & Orgel, L. E. *Science* **219**, 859–862 (1983).
2. Haertle, T. & Orgel, L. E. *J. molec. Evol.* **23**, 108–112 (1986).
3. Zielinski, W. S. & Orgel, L. E. *Nucleic Acids Res.* **13**, 2469–2484 (1985).
4. Zielinski, W. S. & Orgel, L. E. *Nucleic Acids Res.* **15**, 1699–1715 (1987).
5. Von Kiedrowski, G. *Angew. Chem. int. Edn Engl.* **25**, 932–935 (1986).

A possible biochemical missing link among archaeobacteria

Laurie Achenbach-Richter*, Karl O. Stetter†
& Carl R. Woese*

* Department of Microbiology, University of Illinois, Urbana, Illinois 61801, USA

† Department of Microbiology, University of Regensburg, Regensburg, FRG

Until recently all archaeobacteria isolated conformed to one of three basic phenotypes: they were either methanogens, extreme halophiles, or ('sulphur-dependent') extreme thermophiles¹. However, a novel phenotype, that fits none of these categories, has recently been described². The organism, strain VC-16 (tentatively called "*Archaeoglobus fulgidus*") reduces sulphate—the only archaeobacterium so far known to do so—and makes very small quantities of methane, although it lacks some of the cofactors normally associated with methanogenesis². These characteristics suggest that strain VC-16 might represent a transition form between an anaerobic thermophilic sulphur-based type of metabolism (which seems to be the ancestral metabolism for archaeobacteria^{3,4}) and methanogenesis (which somehow then derives from it). We here show that the lineage represented by strain VC-16 arises from the archaeobacterial tree precisely where such an interpretation would predict that it would, between the *Methanococcus* lineage (which is the deepest of the methanogen branchings) and that of *Thermococcus* (the deepest of all branchings on the methanogen side of the tree).

The sequence of the 16S ribosomal RNA gene from "*Archaeoglobus*" strain VC-16 is shown in Fig. 1. A matrix of evolutionary distances derived from an alignment of various archaeobacterial (and outgroup eubacterial) 16S rRNAs is shown in Table 1, and the resulting phylogenetic tree derived from these evolutionary distances is shown in Fig. 2. The particular branching order (apart from strain VC-16) and relative root placement for this tree have been discussed previously³; its topology is independent of whether the analysis is based upon evolutionary distances or minimum change (maximum parsimony), or whether the full sequence alignment or various truncated versions of it (that eliminate the more evolutionarily volatile positions) are used³. Therefore, the topology of Fig. 2 seems to be relatively well established.

Two independent lines of evidence suggest that the ancestral phenotype of the archaeobacteria is most like that seen in the

AUUCUGGUG	AUCCUGCCAG	AGGCCGCGC	UAUCCGCGUC	GGACUAGGCC	50
AUGCGAGUCA	AGGGGCUUGU	AUCCUUCG	GAUGCAAGC	ACCGGCGGAC	100
GGCUCAGUAA	CACUGGACA	ACUGCCUC	GGUGGGGA	UAACCCGGG	150
AAACUGGGC	UAUCCCCA	UAGGGGUA	GUACUGAAU	GUCCAUUC	200
CGAAAGCGCU	UAGCGCCGA	GGAUGGUCU	GCGGCGAAU	AGGUUGUGG	250
UGGGGUAACG	GCCACCAAG	CCGAAGAUC	GUACGGGCA	UGAGAGUGG	300
AGCCCGAGA	UGGACCCUGA	GACACGGUC	CAGGCCUAC	GGGGCGCAG	350
AGGCGGAAA	CCUCCGAAU	GCGGGAACC	GCGACGGGU	CAGCCGGAGU	400
GCUCGCGCAU	CGCGGGGCU	GUCGGGUGC	CUAAAAGCA	CCCCACAGCA	450
AGGGCGGGC	AAGGCGGGU	GCAGCCGCG	CGGUAAUAC	GGCGGCCGA	500
CGGGUAGUCC	UGGUGUAAA	GCUUAAAGC	UAGGUGUAC	CGAAGCUAC	550
UCCUCCGGGA	AAUCUGGCG	CUUAAACG	AGACUGCCG	AGGAUACUG	600
CAGCCUAGG	ACCGGGAGAG	GCCGGGGUA	UCCCGGAGU	AGGGUGGAA	650
UCCUGUAUC	CCGGAGGAC	CACUGUGGC	GAAGCGCCC	GGCUGAACG	700
GGUCCGACG	UGAGGGACGA	AGGCCAGGG	AGCGAACCG	AUAGAUAAC	750
CGGGUAGUCC	UGGUGUAAA	GCUUAAAGC	UAGGUGUAC	CGAAGCUAC	800
AGCUUCCGG	GUGCCGGAG	GAAGCCGUA	AGUCGCGCG	CUGGGGAGU	850
CGCCGCAAG	GCUGAAACU	AAAGGAAU	GCGGGGAGC	ACUACACCG	900
GUGGAGCCG	CGGUUUAUU	GAUUAACG	CCGGGAAGU	UACCGGGGA	950
GACAGCGGA	UGAAGGUCG	GCUGAAGAC	UUACAGACU	AGCUGAGAG	1000
UGGUGCAUG	CCGCGGUCAG	UUCGUACUG	GAAGCAUCC	GUUAGUACG	1050
GCAACGAGC	AGACCCGCG	CCCCAGUUC	CAGCGGUUC	CUUCGGGGA	1100
GCCGGGACA	CGGGGGGAC	UCCGGGGUA	AAGCCGAGG	AAGGUGCGG	1150
CAACGGCAG	UCCGUAUGC	CCGAAUCCG	CGGGUACAC	GCGGGUACG	1200
AUGCGCGGA	CAUUGGUAC	CGACCCGAA	AGGGUAGGU	AAUCCCUAA	1250
ACCGGCGGA	ACCGGGAUC	GAGGGGUGA	ACUCGCCUC	GUGAACCUG	1300
AAUCCGUAU	AAUCGCGCU	CAAAUUGGG	CGGUAUAUC	GUCCUGUCU	1350
CUUGGACACA	CCGCGGUGA	AGCCACCGA	UGGGCCAGG	GGGGCGGGG	1400
UGGCGGAGG	CCACUUCGA	GCCAGGGUC	GCGAGGGGG	GCUAAGUCU	1450
AACAAGGUA	CCGUAAGGA	AUCGCGGCU	GAUACCCUC	CU	1492

Fig. 1 Sequence of the 16S ribosomal RNA gene from "*Archaeoglobus*" strain VC-16.

Methods. The gene was sequenced by standard procedures¹⁴; a 4.5-kb *EcoRI* fragment was cloned into bacteriophage λ gtWES *EcoRI* arms¹⁵ and isolated using *in vitro* ³²P-labelled *Sulfolobus solfataricus* 16S rRNA as a hybridization probe¹⁴; two *HindIII* subfragments of 1.8 and 2.0 kb were then subcloned into bacteriophage M13mp18 (ref. 16) in both orientations, the former originating upstream of the 16S rRNA gene and extending to the *HindIII* site in the gene, the latter originating at this site and extending into the 23S rRNA gene. Sequencing was by the dideoxynucleotide chain termination method^{14,17} on templates produced in phage M13 genome¹⁶, and primers specific for ribosomal RNAs^{14,18}; >90% of the gene was sequenced on both strands.

extreme thermophiles today. Firstly, the extreme thermophile phenotype is the only one of the major archaeobacterial phenotypes found on both major branches of the tree³ (Fig. 2). Secondly, some lineages of extreme thermophiles—those represented by *Thermococcus* and *Pyrodictum* in particular—are more slowly evolving than any of the methanogen or extreme halophile lineages (see Table 1 or Fig. 2). Phylogenetic studies of bacteria indicate a strong correlation between the rate of evolution as measured by relative rRNA sequence distances and the degree of preservation of the ancestral phenotype^{4,5}; the more rapidly evolving lineages tend to lose far more of the

Table 1 Evolutionary distance matrix for various archaeobacterial 16S rRNAs

Species	Evolutionary distance								
	1	2	3	4	5	6	7	8	9
1. " <i>Archaeoglobus</i> " VC-16	—								
2. <i>Tc. celer</i>	16.2	—							
3. <i>Mc. vannielli</i>	22.7	21.0	—						
4. <i>M. formicicum</i>	23.2	21.5	21.8	—					
5. <i>Ms. hungatei</i>	24.2	26.3	28.8	24.6	—				
6. <i>H. volcanii</i>	29.3	28.1	30.1	27.3	25.2	—			
7. <i>P. occultum</i>	22.2	18.4	26.2	25.6	31.4	33.0	—		
8. <i>S. solfataricus</i>	26.9	24.9	29.0	29.6	34.4	36.3	12.3	—	
9. <i>T. tenax</i>	24.7	20.8	29.9	29.0	33.5	34.1	12.1	17.4	—
10. <i>Tt. maritima</i>	38.3	36.1	45.0	44.1	46.8	48.4	36.2	42.1	38.5

The alignment of 16S rRNA sequences has been produced by a standard procedure⁶. In addition to the "*Archaeoglobus*" strain VC-16 sequence it contains the following 16S rRNA sequences: *Thermococcus* (*Tc.*) *celer* (Achenbach-Richter *et al.*, in preparation), *Methanococcus* (*Mc.*) *vannielli*⁷, *Methanobacterium formicicum*⁸, *Methanospirillum* (*Ms.*) *hungatei*⁹, *Halobacterium volcanii*¹⁰, *Pyrodictum occultum* (Achenbach-Richter *et al.*, in preparation), *Sulfolobus solfataricus*¹¹, *Thermoproteus tenax*¹² and the eubacterium *Thermotoga* (*Tt.*) *maritima*⁵. Only those positions in the alignment represented in all archaeobacterial sequences are considered in the analysis. Evolutionary distances have been calculated from per cent sequence differences by the method of Jukes and Cantor¹³.

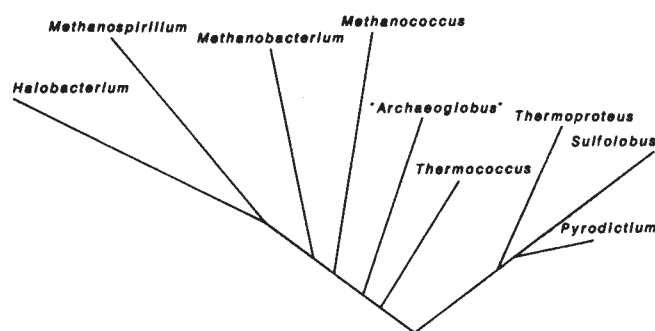


Fig. 2 Phylogenetic tree for representative archaeobacteria, derived from the evolutionary distances of Table 1 by a standard algorithm¹⁹. Line lengths are proportional to evolutionary distances. The root has been determined using the eubacterium *Thermotoga maritima* as an outgroup¹⁴.

ancestral phenotype than do the more slowly evolving lineages. Two independent types of evidence, therefore, argue that all archaeobacterial phenotypes ultimately derive from an ancestral metabolism resembling that seen in the extreme thermophiles today³.

Because its metabolism seems intermediate between methanogenesis and the sulphur reduction of the extreme thermophiles, it is tempting to think of strain VC-16 as a transition form between the extreme thermophiles and methanogens. Like the extreme thermophiles, strain VC-16 is thermophilic and obtains energy from the reduction of sulphur compounds—in this case the more oxidized forms of sulphur, such as sulphate, sulphite, or thiosulphate, however². Also like the extreme thermophiles, the 16S rRNA of strain VC-16 shows higher than random levels of contiguous stretches of G and C residues, which is not the case for methanogens and extreme halophiles (L.A.R. *et al.*, in preparation). On the other hand, the organism produces some methane, though its level is 0.1% or less of that expected for a typical methanogen². Moreover, the organism has several of the cofactors, F420 and methanopterin, associated with methane biosynthesis, although it appears to lack others such as CoM and F430 (ref. 2). When the phylogenetic position of the organism (see Fig. 2) is also taken into account, the possibility that strain VC-16 represents a transition stage in the evolution of (mesophilic) methanogenesis from some thermophilic sulphur-based metabolism becomes worthy of consideration. These studies were supported by grants from the Office of Naval Research and NASA (to C.R.W.) and Deutsche Forschungsgemeinschaft (to K.O.S.). The sequence data in this publication have been submitted to the EMBL/GenBank Libraries under the accession number Y00275.

Received 23 March; accepted 29 April 1987.

- Jones, W. J., Nagle, D. P., Jr & Whitman, W. B. *Microbiol. Rev.* **51**, 135–177 (1987).
- Stetter, K. O., Lauerer, G., Thomm, M. & Neuner, A. *Science* **236**, 822–823 (1987).
- Woese, C. R. & Olsen, G. J. *Syst. appl. Microbiol.* **7**, 161–177 (1986).
- Woese, C. R. *Microbiol. Rev.* **51**, 221–271 (1987).
- Woese, C. R., Stackebrandt, E. & Ludwig, W. *Molec. Evol.* **21**, 305–316 (1985).
- Woese, C. R., Gutell, R., Gupta, R. & Noller, H. F. *Microbiol. Rev.* **47**, 621–669 (1983).
- Jarsch, M. & Böcke, A. *Syst. appl. Microbiol.* **6**, 54–59 (1985).
- Lechner, K., Wich, G. & Böck, A. *Syst. appl. Microbiol.* **6**, 157–163 (1985).
- Yang, D., Kaine, B. P. & Woese, C. R. *Syst. appl. Microbiol.* **6**, 251–256 (1985).
- Gupta, R., Lanter, J. & Woese, C. R. *Science* **221**, 56–659 (1983).
- Olsen, G. J. *et al. J. molec. Evol.* **22**, 301–307 (1985).
- Leinfelder, W., Jarsch, M. & Böck, A. *Syst. appl. Microbiol.* **6**, 164–170 (1985).
- Jukes, T. H. & Cantor, C. R. in *Mammalian Protein Metabolism* (ed. Munro, H. N.) 21–132 (Academic, New York, 1969).
- Achenbach-Richter, L., Gupta, R., Stetter, K. O. & Woese, C. R. *Syst. appl. Microbiol.* **9**, 34–39 (1987).
- Leder, P., Tiemeier, D. & Enquist, L. *Science* **196**, 175–177 (1977).
- Messing, J. *Meth. Enzym.* **101**, 20–78 (1983).
- Biggin, M. D., Gibson, T. J. & Hong, G. F. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3963–3965 (1983).
- Yang, D. C., Kaine, B. & Woese, C. R. *Syst. appl. Microbiol.* **6**, 251–256 (1985).
- De Soete, G. *Psychometrika* **48**, 621–626 (1983).

High resolution X-ray analyses of renin inhibitor-aspartic proteinase complexes

S. I. Foundling*, J. Cooper*, F. E. Watson*, A. Cleasby*, L. H. Pearl*, B. L. Sibanda*, A. Hemmings*, S. P. Wood*, T. L. Blundell**, M. J. Valler†, C. G. Norey†, J. Kay†, J. Boger‡, B. M. Dunn§, B. J. Leckie||, D. M. Jones||, B. Atrash¶, A. Hallett¶ & M. Szelke¶

* Laboratory of Molecular Biology, Department of Crystallography, Birkbeck College, Malet Street, London WC1E 7HX, UK

† Department of Biochemistry, University College, PO Box 78, Cardiff CF1 1XL, UK

‡ Merck Sharpe and Dohme Research Laboratories, PO Box 2000, Rahway, New Jersey 07065, USA

§ Department of Biochemistry and Molecular Biology, University of Florida College of Medicine, Gainesville, Florida 32610, USA

|| MRC Blood Pressure Unit, Western Infirmary, Glasgow G11 6NT, UK

¶ Department of Chemical Pathology, Royal Postgraduate Medical School, London W12 0HS, UK

Inhibitors of the conversion of angiotensinogen to the vasoconstrictor angiotensin II have considerable value as antihypertensive agents. For example, captopril and enalapril are clinically useful as inhibitors of angiotensin-converting enzyme. This has encouraged intense activity in the development of inhibitors of kidney renin, which is a very specific aspartic proteinase catalysing the first and rate limiting step in the conversion of angiotensinogen to angiotensin II. The most effective inhibitors such as H-142 (ref. 1) and L-363,564 (ref. 2) have used non-hydrolysable analogues of the proposed transition state^{3–6}, and partial sequences of angiotensinogen (Table 1)^{7–9}. H-142 is effective in lowering blood pressure in humans^{10,11} but has no significant effect on other aspartic proteinases such as pepsin in the human body (Table 1). At present there are no crystal structures available for human or mouse renins¹² although three-dimensional models^{13–16} demonstrate close structural similarity to other spartic proteinases. We have therefore determined by X-ray analysis the three-dimensional structures of H-142 and L-363,564 complexed with the aspartic proteinase endothiapepsin^{17–19}, which binds these inhibitors with affinities not greatly different from those measured against human renin (Table 1). The structures of these complexes and of that between endothiapepsin and the general aspartic proteinase inhibitor, H-256 (Table 1) define the common hydrogen bonding schemes that allow subtle differences in side-chain orientations and in the positions of the transition state analogues with respect to the active-site aspartates.

X-ray data were collected on crystals isomorphous with the native for complexes of endothiapepsin with H-142, L-363,564 and H-256 to resolutions of 2.1, 2.2 and 2.0 Å and refined by least-squares to agreement values of 0.19, 0.18 and 0.20 respectively. Figure 1 shows the inhibitors that lie with extended conformations in the deep cleft between the NH₂- and COOH-terminal lobes of the aspartic proteinase^{17,19}. The putative transition state isosteres at P₁-P_{1'} are in close juxtaposition with the two catalytic carboxyl groups of Asp 32 and Asp 215 (pepsin numbering) which are in the centre of the cleft^{17,20,21}. In general, the main-chain torsion angles define β -strands P₄-P₁ and P_{1'}-P₃. The residues in P₃ to P₁ of each inhibitor form an antiparallel pleated sheet with residues 219 to 217 of the enzyme (Fig. 2). The hydrogen bonding arrangement is extended as far as the NH group of the P₃ residue which is hydrogen bonded to the γ -O of Thr 219 as described previously in complexes of pepstatin (fragments) with penicillopepsin and rhizopuspepsin^{22,23}.

* To whom correspondence should be addressed.

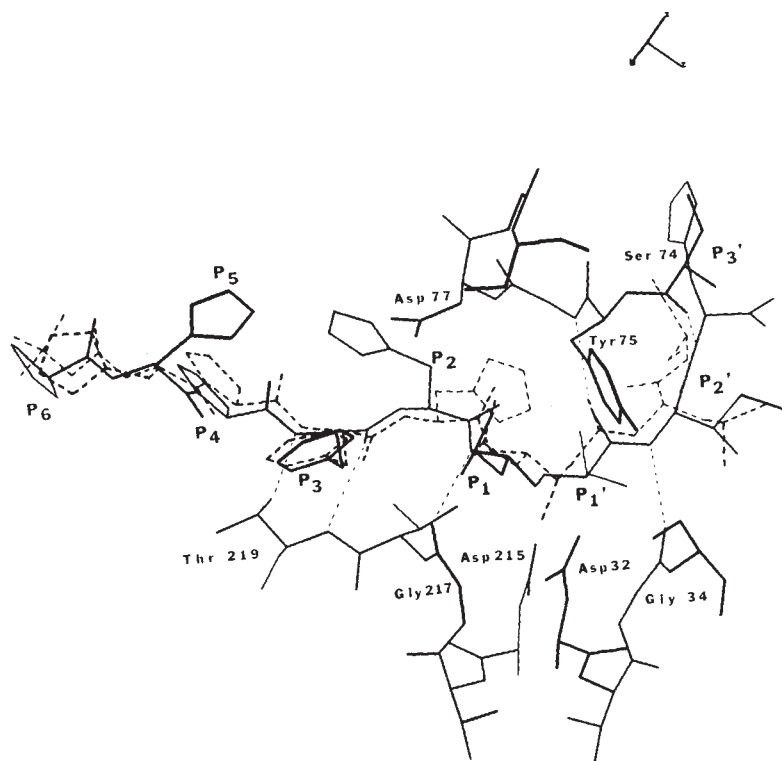


Fig. 1 Views of H142 (solid line) and L-363,564 (broken line) superposed as they are bound to the active site of endothiapsin. X-ray data for the various inhibitors were collected on Hilger and Watts Y290 and Nonius CAD4 diffractometers and corrections were applied for absorption, Lorentz, polarization and radiation decay factors. The R factors for symmetry-related reflections are 0.13, 0.06 and 0.06 for H-142, L-363,564 and H-256 respectively. Electron density maps were first calculated with weighted coefficients ($F_{pd} - F_p$) and ($2F_{pd} - F_p$) and phases derived from the refined structure of endothiapsin. They were displayed on an Evans and Sutherland PS2 using the program FRODO²⁷ modified by I. J. Tickle. The two electron density maps were used to define the conformation of the inhibitor and to suggest changes in the conformation of the enzyme and positions of the water molecules. Cycles of restrained least squares refinement of all atomic coordinates and individual temperature factors (B_{iso}) using the program RESTRAIN²⁸ and model building using electron density maps calculated using the refined phases for the inhibitor complex led to a structure with good geometry and acceptable agreement values.

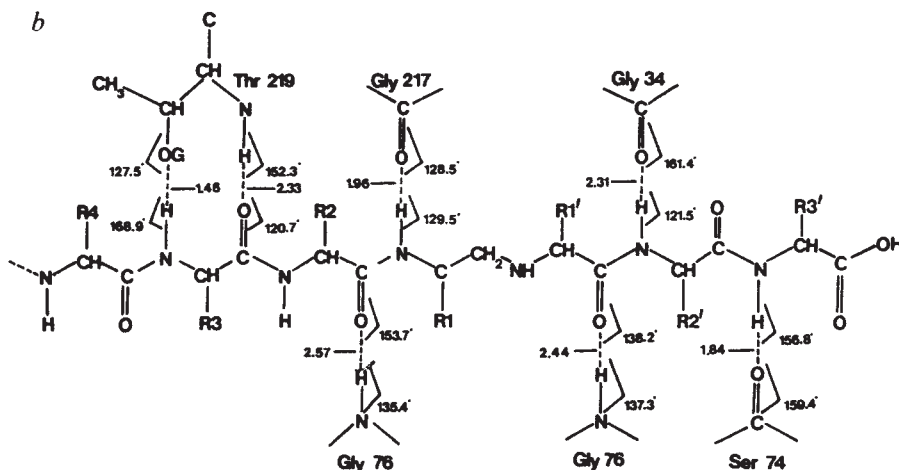
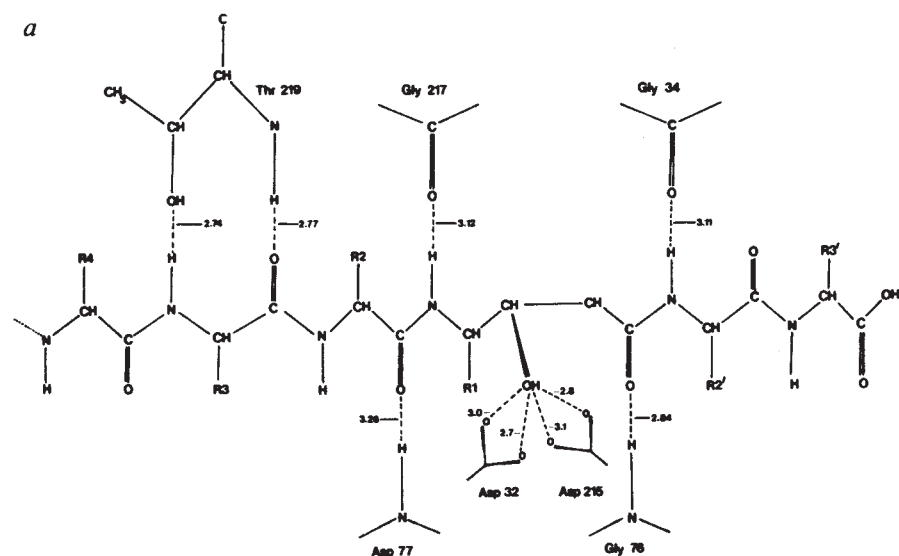


Fig. 2 A schematic representation of the hydrogen bonding between the enzyme and *a*, L-363,564 or *b*, H-142. Both inhibitors displace the water or cation hydrogen bonded to the carboxylic acid groups of Asp 32 and Asp 215 in the native enzyme.

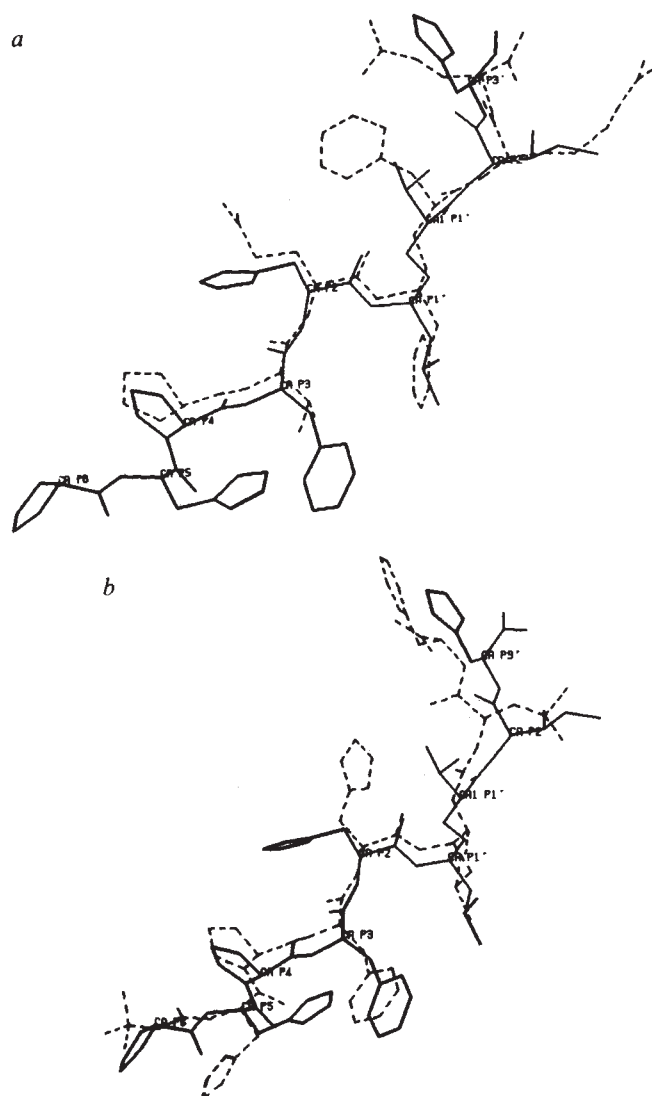


Fig. 3 The conformations of three inhibitors shown superposed. *a*, H-142 (solid lines) and H-256 (broken lines); *b*, H-142 (solid lines) and L-363,564 (broken lines).

Residue 219 is maintained as either threonine or serine in all aspartic proteinases, so that interactions with this residue may be important to the alignment of the substrate/inhibitor in the active site. The topologically equivalent residues, 34 to 36 (related by a pseudo dyad to 217 to 219 of COOH-terminal domain as described by Tang *et al.*²¹) do not appear to participate in a corresponding parallel β -sheet interaction with the P_1' - P_3' residues. Only the NH-group of the P_2' residue of the inhibitors makes such an interaction (Fig. 2), giving a good hydrogen bond with Gly 34 CO-group of the enzyme; P_3' is oriented towards the solvent.

Although the orientations of the main chain and side chains of the three inhibitors are similar at the well-defined residues P_4 and P_3 , that bind in shallow pockets, the more exposed side chains of P_6 , P_5 and P_2 show considerable variation in their interactions (Fig. 3). For example, the side-chain imidazole of the histidine in P_2 is oriented in two different ways (Fig. 3). In L-363,564, the P_2 imidazole is oriented more towards the interior than in H-142 and makes contacts with residues that participate in S_1' interactions in the H-142 complex. This distinction may be due, in part, to the absence of a P_1' residue in L-363,564, which contains the statine residue as a dipeptide analogue. However, the P_2 Glu residue in H-256 adopts an intermediate position interacting with solvent. These observations suggest

Table 1 Kinetic constants (K_i) for the inhibition of aspartic proteinases by three synthetic inhibitors

Enzyme	Inhibitor		
	L-363,564	H-142	H-256
Human pepsin	130	>40,000	30
Human gastricsin	200	15,000	60
Human cathepsin D	110	>40,000	150
Human renin	3	10	>7,600
Penicillopepsin	300	24,000	10
Endothiapepsin	16	160	60

Values are K_i in nM. All assays were initiated after 5 min preincubation of inhibitor and enzyme at 37 °C in 0.1 M sodium formate buffer, pH 3.1 by the addition of a synthetic chromogenic peptide substrate, Lys-Pro-Xaa-Yaa-Phe-Phe(4-NO₂)-Arg-Leu (ref. 26). Human renin was assayed at 37 °C using angiotensinogen as substrate in 0.1 M sodium citrate/phosphate buffer, pH 7.2. The angiotensin I released was measured by radio-immunoassay¹. The angiotensinogen and inhibitor sequences are:

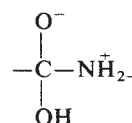
Angiotensinogen	6	8	10	12	
H-142	~His	Pro	Phe	His	Leu
L-363,564	Pro	His	Pro	Phe	His
H-256	Boc	His	Pro	Phe	His
	Pro	Thr	Glu	Phe	His
	P_6	P_5	P_4	P_3	P_2
					P_1
					P_1'
					P_2'
					P_3'

The statine-containing compound L-363,564 (Table 1) was synthesized as described by Boger *et al.*². H-142, containing the reduced ($-\text{CH}_2\text{NH}-$) isostere of the scissile ($-\text{CONH}-$) peptide bond between residues Leu 10 and Val 11 in human angiotensinogen, was obtained as in Szelke *et al.*¹. The peptide, H-256 (Table 1) was synthesised as described¹¹ to contain the reduced isostere of a Phe-Phe dipeptide and to be sufficiently long to occupy the seven recognised sub-sites (S_4 to S_3').

that it would be unwise to establish rigid definitions of enzyme subsites on the basis of one inhibitor. Inhibitors—and possibly substrates—tend to be rather promiscuous, exploring relationships with various residues of the enzyme in their vicinity.

The hydrophobic sidechains of leucine, phenylalanine or statine at P_1 occupy a deep pocket which is an extension of S_3 and involves the hydrophobic side chains of Tyr 75, Phe 111 and Leu 120. In the H-142 and H-256 complexes, the hydrophobic side chains of P_1' interact with a similar topologically related hydrophobic pocket comprising Ile 297, Ile 301 and Ile 299. The bulkier aromatic side chain of the Phe at P_1' of H-256 appears to be inserted more deeply into the pocket, probably as a result of more extensive hydrophobic interactions. In the inhibitor L-363,564, the statine residue bridges both P_1 and P_1' positions, but has no side chain to correspond to that of a substrate in the P_1' position.

The similar hydrophobic bonding interactions and positions of P_4 , P_3 and P_1 in hydrophobic pockets locate the transition state analogues of the scissile peptide bond close to the carboxylates of the two active aspartic acid residues, 32 and 215. In L-363,564, the hydroxyl of the statine residue displaces the water molecule or cation and occupies its position in the active site of the native enzyme (Fig. 2). The symmetry of bonding of this hydroxyl to the carboxylates of both Asp 32 and Asp 215 is apparent. Thus, during the hydrolysis of a peptide substrate, it would seem that one of the oxygens of the tetrahedral intermediate



is likely to occupy this position although from this evidence it is not possible to distinguish whether the oxygen is derived from the carbonyl oxygen of the peptide bond or the water molecule attacking it. Neither H-142 nor H-256 contain a hydroxyl group but both have a secondary amine. In the H-142 structure, the

main chain has flexed to allow the nitrogen atom to approach the plane of the two carboxylates in a similar way to that of the ϵ -amino group of a lysine residue in the pepsinogen structure²⁴. This might be interpreted to mean that it is the amino group which is protonated and which interacts with the carboxylates in the transition state. In the H-256 structure, the amino group is positioned less close to the two carboxylates (Fig. 3) apparently as a result of the deeper penetration of the Phe P₁' into the S₁' subsite in the H-256 complex as described above.

Few changes of conformation can be detected between the native enzyme structure and the protein in the inhibited complexes. The exception is a slight repositioning of the residues in the 'flap' from the average position that they occupy in the native enzyme. This region is highly flexible in the native enzyme and must open to allow access of the substrate or inhibitor to the active site cleft. In the complexes, the flap covers the catalytic groups of the enzyme and P₁-P₁' region of the inhibitor to give yet further hydrogen bonds. The flap thus ensures an environment generally secluded from the solvent so that transient intermediates are encapsulated by the enzyme. Small movements do occur in the region of S₃ but these are not propagated extensively through the enzyme. The major change resulting from inhibitor binding is a lowering of the thermal and static disorder in the cleft residues of the complexes as indicated by significant decreases (between 30 and 50%) in the thermal factors, B_{iso} . This would be expected of the van der Waals, ionic and hydrogen bonds made between the enzyme and the inhibitor.

The X-ray studies described here define possible conformations of binding of inhibitors which are likely to be relevant to the homologous enzyme, human renin. They provide the basis for detailed modelling to test this hypothesis, and may also provide clues for the synthesis of analogues which are effective inhibitors of human renin, as a result of a better complementarity of shape, charge and hydrophobicity in all the sub-sites. Intramolecular cyclization to increase rigidity and stabilize the bound conformation of the transition state analogues should also improve their effectiveness²⁵.

The Birkbeck and Cardiff groups thank the SERC and the MRC, and the RPMS group thank Ferring, for financial support. We thank Dr Bruce Bush (Merck, Sharpe and Dohme) and Dr Ian Tickle (Birkbeck) for useful discussions.

Received 16 March; accepted 7 April 1987.

- Szelke, M. *et al.* *Nature* **299**, 555-557 (1982).
- Boger, J. *et al.* *Nature* **303**, 81-84 (1983).
- Hallett, A. *et al.* in *Aspartic Proteinases and their Inhibitors* (ed. Kostka, V.) 467-478 (de Gruyter, Berlin, 1985).
- Dunn, B. M., Kammermann, B. & McCurry, K. *Analyt. Biochem.* **138**, 68-73 (1984).
- Szelke, M., Jones, D. M., Atrash, B., Hallett, A. & Leckie, B. J. in *Peptides, Structure and Function: Proceedings of the Eighth American Peptide Symposium* (eds Hruby, V. J. & Rich, D. H.) 579-582 (Pierce Chemical, Rockford, Illinois, 1983).
- Szelke, M. in *Aspartic Proteinases and their Inhibitors* (ed. Kostka, V.) 421-441 (de Gruyter, Berlin, 1985).
- Skeggs, L. T., Lentz, K., Hochstrasser, H. & Kahn, J. R. *Can. med. Ass. J.* **90**, 185-189 (1964).
- Kokusbu, T. *et al.* *Nature* **217**, 456-457 (1968).
- Haber, E. & Burton, J. *Fedn Proc.* **38**, 2768-2773 (1979).
- Webb, D. J. *et al.* *Lancet* **ii**, 1486-1487 (1983).
- Webb, D. J. *et al.* *Hypertension* **3**, 653-658 (1985).
- Navia, M. A., Springer, J. P., Poe, M., Boger, J. & Hoogsteen, K. *J. biol. Chem.* **259**, 12714-12717 (1984).
- Blundell, T. L., Sibanda, B. L. & Pearl, L. H. *Nature* **304**, 273-275 (1983).
- Sibanda, B. L. *et al.* *FEBS Lett.* **174**, 102-111 (1984).
- Akahane, K. *et al.* *Hypertension* **7**, 3-12 (1985).
- Carlson, W., Karplus, M. & Haber, E. *Hypertension* **7**, 13-19 (1985).
- Pearl, L. H. & Blundell, T. L. *FEBS Lett.* **174**, 96-101 (1984).
- Pedersen, V. B. *Biochemistry* (in the press).
- Blundell, T. L., Jenkins, J., Pearl, L. H., Sewell, T. & Pedersen, V. in *Aspartic Proteinases and their Inhibitors* (ed. Kostka, V.) 151-161 (de Gruyter, Berlin, 1985).
- James, M. N. G. & Sielecki, A. *J. molec. Biol.* **163**, 299-361 (1983).
- Tang, J., James, M. N. G., Hsu, I. N., Jenkins, J. A. & Blundell, T. L. *Nature* **271**, 618-621 (1978).
- Bott, R., Subramanian, E. & Davies, D. R. *Biochemistry* **21**, 6956-6962 (1982).
- James, M. N. G., Sielecki, A. R., Salituro, F., Rich, D. H. & Hofmann, T. *Proc. natn. Acad. Sc. U.S.A.* **79**, 6137-6141 (1982).
- James, M. N. G. & Sielecki, A. R. *Nature* **319**, 33-39 (1986).
- Boger, J. in *Aspartic Proteinases and their Inhibitors* (ed. Kostka, V.) 401-420 (de Gruyter, Berlin, 1985).
- Dunn, B. M. *et al.* in *Aspartic Proteinases and their Inhibitors* (ed. Kostka, V.) 221-243 (de Gruyter, Berlin, 1985).
- Jones, T. A. *Appl. crystallogr.* **11**, 268-272 (1978).
- Haneef, I., Moss, D. S., Stanford, M. J. & Borkakoti, N. *Acta crystallogr.* **A41**, 426-433 (1985).

Copies of articles from this publication are now available from the UMI Article Clearinghouse.

For more information about the Clearinghouse, please fill out and mail back the coupon below.

UMI Article Clearinghouse

Yes! I would like to know more about UMI Article Clearinghouse. I am interested in electronic ordering through the following system(s):

- ☐ DIALOG/Dialorder ☐ ITT Dialcom
☐ OnTyme ☐ OCLC ILL Subsystem
☐ Other (please specify) _____
☐ I am interested in sending my order by mail.
☐ Please send me your current catalog and user instructions for the system(s) I checked above.

Name _____

Title _____

Institution/Company _____

Department _____

Address _____

City _____ State _____ Zip _____

Phone (_____) _____

Mail to: University Microfilms International
300 North Zeeb Road, Box 91 Ann Arbor, MI 48106